

A Fully Optical Ring Network-on-Chip with Static and Dynamic Wavelength Allocation

CISSE AHMADOU DIT ADI



Graduate School of Information Systems

Department of Network Systems

University of Electro-Communications, Tokyo, Japan

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Engineering

MARCH 2014

A Fully Optical Ring Network-on-Chip with Static and Dynamic Wavelength Allocation

Approved by Supervisory Committee

Chair Prof. Tsutomu Yoshinaga

Member Prof. Hiroshi Nagaoka

Member Prof. Maomi Ueno

Member Prof. Yoshikatsu Tada

Member Asso. Prof. Satoshi Ohzahata

©CISSE AHMADOU DIT ADI

Japanese Abstract

プロセッサやメモリなど、単一の半導体チップに集積されるIP(Intellectual Property)コアが増大するにつれて、Chip Multicore Processor(CMP)やSystem-on-Chips(SoC)における高性能かつ低消費電力な相互接続基盤が重要な役割を果たすようになってきている。帯域の制限やクロストーク問題、インピーダンス不整合、巨大なエネルギー損失など、従来の電子的な接続網によるNetwork-on-Chip(NoC)は様々な問題に直面している。これらの困難を軽減する有望な解決手法として、光NoCが注目されている。光インターコネクは、単一のオプティカルリンク(waveguide)における波長の多重化を利用した光通信により、低い消費電力で高通信帯域を実現する。本研究は、以下に示す3つの研究を通して、高いコスト性能比と電力効率を持つNoCを提案する。

(1) 伝統的な電子-光トラス型のNoCにおいて、予測スイッチングを用いて低レイテンシな経路を確立するネットワークを提案する。経路確立レイテンシの軽減により、大きな性能向上を得ることができる。

(2) 光リングと電子的なクロスバによって構成されるハイブリッドアーキテクチャOREX(Optical-Ring and Electrical-Crossbar)を提案する。OREXは、単一のクロスバに比べて経路確立時間を低減する。OREXの光ネットワークは、光インターコネクにより適したリングトポロジを構成する。サイクル精度のシミュレータを用いて、OREXが光通信と電子通信のハイブリッドNoCの性能をより向上されることを示す。

(3) より消費電力を低減するため、完全光リングアーキテクチャを提案する。提案アーキテクチャは、低消費電力かつ高性能な光インターコネクを利用して、同じネットワークに対する静的および動的な波長の割り当て手法を統合する。軽量の通信は、異なる波長チャネルが各宛先ノードに静的に割り当てられる。複数の送信元ノードから同一の宛先ノードへの通信リクエストの同時発生による競合は、トークンによる調停が行われる。高負荷な通信においては、送信元ノードから動的な波長割り当てを管理する特殊なノードに実行時間が要求され、共有された多重波長チャネルが利用される。本研究を通して提案するアーキテクチャは、通信メッセージサイズ(baseline)にしたがって適切な波長を選択する波長割り当て手法と、ネットワークの混雑情報(競合と高性能な選択)を利用する。複数の光アーキテクチャとの予備的なハードウェアコスト比較を通して、提案アーキテクチャは将来のSoCやCMPにおける相互接続基盤として有望なコストパフォーマンスが得られることを示す。さらに、光ネットワークシミュレータを用いたシミュレーションを通して、提案した完全光リング型NoCアーキテクチャの性能について議論する。提案アーキテクチャは従来型のハイブリッドNoCと比較して、消費エネルギーの大きな削減が可能であることを示すとともに、確率的なトラフィックパターンを用いて実用的な帯域とレイテンシが得られることを示す。

Abstract

As the number of IPs (processor, memory) integration on a chip increases, chip multicore processors (CMPs), and system-on-chips (SoCs) will require high performance and low power consumption interconnection infrastructure. Traditional electronic network-on-chip (NoC) faces several problems, such as limited bandwidth, crosstalk, impedance mismatch, and huge power dissipation. To alleviate these challenges, optical NoCs have emerged as an attractive solution. Optical interconnects take advantage of light, and the multiple wavelengths within a single optical link (waveguide) to achieve high communication bandwidth at low power consumption cost. Toward this work we aim to propose a cost-performance and power efficient NoC. First, we proposed a low latency path setup network for conventional hybrid electronic-photonic Torus NoC using predictive switching. By lowering the path setup latency, we could achieve a considerable performance improvement. Second, a new hybrid architecture formed of an optical ring and electrical crossbar (OREX) has been proposed. OREX reduces the path setup network to a single electrical crossbar. Its optical network uses a ring topology more adapted for photonic interconnects. Using a cycle accurate simulator, our results show that OREX further improves hybrid electronic-photonic NoCs performance. Finally, to reduce power consumption, we have proposed a fully optical ring architecture. The proposed architecture combines static and dynamic wavelength allocation in the same network to fully take advantage of the low power and high performance optical interconnects. A different wavelength-channel is statically allocated to each destination node for light weight communication. Contention of simultaneous communication requests from multiple source nodes to the destination is solved by a token based arbitration. For heavy load communication, a shared multiwavelength-channel is available by requesting it in execution time from source node to a special node that manages dynamic wavelength allocation. Our architecture takes advantage of both wavelength allocation mechanisms by selecting the adequate one, depending on

communication message sizes (baseline) and network congestion information (contention based and smart selection). Preliminary hardware cost comparison with several photonic architecture shows that our architecture can be an attractive cost-performance interconnection infrastructure for future SoCs and CMPs. We further discuss performance of the proposed fully optical ring NoC architecture based on simulation using a photonic network simulator. Results show that our architecture allows considerable reduction of the network energy consumption compared to conventional hybrid NoCs and show reasonable bandwidth and latency performance using probabilistic traffic patterns.

CONTENTS

Abstract	iv
Acknowledgments	xiv
Abbreviations	xv
1 Introduction	1
1.1 Problem Definition	2
1.2 Approach and Contributions	2
1.3 Thesis Overview	4
2 Overview of on-Chip Optical interconnect and Related Works	5
2.1 Optical Devices	5
2.1.1 Light Source	6
2.1.2 Waveguide	7
2.1.3 Modulator	7
2.1.4 Detector	8
2.2 Related Works	8

CONTENTS

2.2.1	Hybrid Circuit-Switch	9
2.2.2	CORONA	9
2.2.3	Firefly	10
2.2.4	FlexiShare	10
2.2.5	PROPEL	11
2.2.6	All Optical Routed Wavelength Architecture	11
3	Low latency Path Setup Hybrid Photonic Torus	13
3.1	Hybrid Photonic Torus	13
3.1.1	Optical Network	16
3.1.2	Electrical Network	18
3.1.2.1	Predictive Switching Based Path Setup	18
3.1.2.2	Reservation Based Path Setup	22
3.2	Performance Evaluation	24
3.2.1	Power Consumption Estimation	24
3.2.2	Simulation Conditions	26
3.2.3	Results and Discussion	28
3.3	Conclusion	34
4	OREX: Hybrid Optical Ring Electrical Crossbar Network-on-Chip	35
4.1	OREX Architecture	36
4.2	Communication Mechanism	39
4.2.1	Path Setup and Optical Data Transfer	39
4.2.2	Path Release	41
4.3	Wavelength Allocation	42
4.3.1	Single Data Stream Wavelength Allocation	42
4.3.2	Multiple Data Stream Wavelength Allocation	42
4.4	Cost Comparison	43
4.4.1	Hardware cost	43

CONTENTS

4.4.2	Power Consumption	44
4.4.3	Bandwidth	45
4.5	Performance Evaluation	47
4.5.1	Simulation Conditions	47
4.5.2	Experimental Results	50
4.5.2.1	Zero-load latency	50
4.5.2.2	Latency Evaluation under uniform and neighbor traf- fic patterns	50
4.6	Conclusion	54
5	Fully Optical Ring Network-on-Chip	55
5.1	A Fully Optical Ring NoC's Architecture	56
5.2	Communication Mechanisms	57
5.2.1	Static Communication	58
5.2.2	Dynamic Communication	59
5.2.3	Bended Static and Dynamic Communications	61
5.3	Wavelength Allocation Selection Mechanisms	62
5.3.1	Baseline Selection Mechanism	62
5.3.2	Contention Based Selection Mechanism	63
5.3.3	Smart Selection Mechanism	64
6	Simulation Results and Analysis	66
6.1	Hardware Cost Comparison	66
6.2	Simulation Setup	67
6.3	Static and Dynamic Communications Comparison	71
6.4	Performance and Energy Consumption Comparison	72
6.5	FORNoC with different selection techniques	75
6.6	FORNoC under Partially Localized and Localized Probabilistic Traf- fic Patterns	78

CONTENTS

6.7 Scalability	81
6.8 Conclusion	82
7 Summary	83
7.1 Conclusion	83
7.2 Future works	84
Publications	85
Bibliography	87

LIST OF TABLES

2.1	Related works	12
3.1	Advantages and disadvantages of CPS and RPS	24
3.2	Simulation parameters	27
3.3	Optical NoC parameters	27
3.4	Electrical NoC parameters	27
4.1	Hardware cost comparison	44
4.2	Simulation parameters	47
4.3	OREX simulations configurations	49
6.1	Architecture hardware cost comparison for 64-node networks.	68
6.2	Simulation parameters	70

LIST OF FIGURES

2.1	Block diagram of on-chip optical interconnect.	6
3.1	A 4×4 hybrid torus photonic NoC	15
3.2	Optical switch [7]	17
3.3	Electrical routers	19
3.4	Pipeline time diagram for conventional and prediction routers	20
3.5	Example of prediction using SPM scheme	21
3.6	(a) Conventional vs (b) reservation based path setup mechanisms . .	22
3.7	Power consumption cost	24
3.8	Electrical NoC under uniform (a), neighbor (b), and bitreversal traffic patterns (c)	28
3.9	HPNoC under uniform (a), neighbor (b), and bitreversal traffic pat- terns (c)	30
3.10	Electrical NoC vs HPNoC under uniform (a), neighbor (b), and bitre- versal traffic patterns (c)	32
3.11	HPNoC, CPS vs RPS under uniform (a), and bitreversal traffic pat- terns (b)	33

LIST OF FIGURES

4.1	An 8 nodes OREX topology	36
4.2	Connection between a node and routers	38
4.3	Path setup (1, 2, 3), and optical data transfer (4) sequences of OREX communication.	39
4.4	Optical path release processes	41
4.5	Single and two data streams waveguides	43
4.6	Power consumption comparison	45
4.7	Bandwidth comparison	46
4.8	A time diagram of communication on OREX.	50
4.9	Zero load latency.	51
4.10	Average latency of OREX under uniform (a, c), and neighbor(b, d) traffic patterns.	52
4.11	Effect of path multiplicity using single and multiple waveguides under uniform (a, c) and neighbor (b, d) traffic patterns.	53
5.1	FORNoC architecture	57
5.2	Nodes microarchitecture.	58
5.3	Nodes connection in the static communication waveguide.	59
5.4	Time diagram of a static communication	60
5.5	Time diagram of a dynamic communication	61
5.6	Smart selection mechanism.	65
6.1	Performance comparison of static and dynamic communications	71
6.2	Energy consumption comparison.	73
6.3	Latency and bandwidth performance.	74
6.4	Low load latency under uniform traffic (SP)	76
6.5	Latency performance comparison of FORNoCs.	77
6.6	Neighbor Traffic pattern.	79
6.7	Hotspot traffic pattern.	80

LIST OF FIGURES

6.8 Average latencies for 32, 64, and 128-node under uniform random traffic.	81
--	----

Acknowledgments

Bismillahi Arahamani Arahim

This Doctoral project has been accomplished at the laboratory of Network Computing in the Department of Information Network Systems of the Graduate school of Information Systems, in the University of Electro-communication(UEC). Many people have contributed in one way or another for the accomplishment of This work.

I am very grateful to *ALLAH AZWAJAL* for guaranteed me health to be able to accomplish this work.

Thanks to My mother, father, brothers, relatives and friends for their kind support. A special thanks to my supervisor Prof. Yoshinaga Tsutomu for his instructions and supports in all stages of this work.

I would like to also thanks Michihiro Koibuchi, Hidetsugu Irie and Masato Yoshimi for their kind helps and advices.

My thanks goes also to the Laboratory members for their daily encouragement.

Finally many thanks to all the professors and office staff members of UEC.

Abbreviations

CMP	Chip Multicore Processors
CMOS	Complementary Metal-Oxide Semiconductor
CPS	Conventional Path Setup
DOR	Dimension Ordered Routing
ENoC	Electrical Network-on-Chip
EO	Electrical to Optical conversion
FORNoC	Fully optical Ring Network-on-Chip
HPNoC	Hybrid Photonic Network-on-Chip
HPTNoC	Hybrid Photonic Torus Network-on-Chip
LAN	Local Area Network
LP	Last Port
LT	Link Traversal
MR	Micro-ring Resonator
MWSR	Multiple Write single Read
NI	Network Interface
NoC	Network-on-Chip
OA	Optical Allocation
OE	Optical to Electrical conversion
OREX	Optical Ring Electrical Crossbar
OS	Optical Switch setting
OT	Optical Traversal
PMNoC	Planar Mesh Network-on-Chip
RC	Routing Computation
RPS	Reservation Based Path Setup
RR	Read Request

SAN	Storage Area Network
SoC	System-on-Chip
SPM	Sample Pattern matching
SWMR	Single Write Multiple Read
TG	Token Grant
TR	Token Release
VSA	Virtual Channel and Switch Allocation
WA	Wavelength Allocation
WAN	Wide Area Network
WDM	Wave Division Multiplexing

CHAPTER 1

INTRODUCTION

Transistor size is continuously shrinking down, leading to better chip integration capabilities. According to the international technology roadmap for semiconductors (ITRS), hundreds of cores can be integrated in a single chip in near future. Therefore, the communication infrastructure should be improved to deal with the enormous increases in complexity, energy consumption, and bandwidth demands. Today's electrical network-on-chips (NoCs), which consume a huge amount of power for electrical signaling, face critical challenges to provide the required communication performance within the available power budget. These limitations direct current research activities on finding alternative approaches with better performance and energy efficiency.

After their performance and power efficiency capabilities have been proven for many applications, ranging from wide area networks (WANs), Local area Networks (LANs), and storage area networks(SANs), Optical interconnects are making their way to the chip level. With Recent development of CMOS compatible optical devices, their usage at the chip level are becoming more realistic. Optical intercon-

nects have the intrinsic capabilities to provide high data transfer throughput and low latency at low power consumption cost compare to their electrical counterpart. In this thesis we will focus on proposing ways to use current available optical devices to provide a better cost, performance and energy efficient interconnection infrastructure for multicore processors (CMPs) and system-on-chips (SoCs).

1.1 Problem Definition

As the number of cores on a chip increases, many-core and system-on-chips (SoCs) interconnections will require high performance and low power consumption. Traditional electronic network-on-chip (NoC) faces several problems, such as limited bandwidth, crosstalk, impedance mismatch, and huge power dissipation. Photonic communication technology offers an opportunity to reduce the interconnection power consumption while meeting future chip multiprocessors (CMPs) performance requirements. It has attracted attention with recent advances on development of required silicon photonics devices. CMOS-compatible micro-ring resonators (MRs [21, 29]), photonic detectors [22], and silicon waveguides [9, 5], key devices which are necessary to integrate photonic network at the chip level. Several works that combine photonic and electronic interconnects (hybrid NoCs [12, 25, 13, 18, 20, 4]) or use pure optical interconnects (fully optical NoCs [2, 28, 14, 27]) showed that silicon photonics could be a promising solution for future NoCs.

1.2 Approach and Contributions

Toward this work our goal is to propose ways to use current optical interconnects to achieve a cost-performance and power efficient NoC. With optical interconnects lacking optical data processing and buffering, to take advantage of optical interconnects, the first integration of optical interconnects inside a chip suggested a hybrid

photonic network-on-chip architecture(HPNoC) [25]. The architecture consists of a photonic layer, which uses a high-bandwidth circuit switching, controlled by an electrical packet switching layer. The HPNoC removes the need for buffering of optical data and the high power consumption of optical-electrical-optical (O-E-O) conversions at intermediate node for routing computation. With the combination of the optical circuit-switching network and electrical packet-switching network, the HPNoC provides a better interconnection bandwidth and transmission speed at a lower power consumption in comparison with an all-electrical NoC architecture [23]. However the performance of such architecture depend in large on how fast optical paths are set for communication.

Because it is critical for the electrical NoC of a HPNoC fabrics to have low latency, First we proposed a low latency electrical control network for such HPNoC architecture using predictive switching and reservation based path setup techniques. Predictive switching speculatively forward the packets inside a router bypassing some pipeline stages. It allows a considerable performance improvement other conventional switching techniques. The reservation based path setup technique reserves path ahead a time to also reduce the path setup latency of the hybrid architecture allowing an overall improvement of network performance.

Second, we further improve the performance of HPNoC by proposing a new interconnection architecture formed of an optical ring and electrical crossbar (OREX). OREX reduces the path setup network to a single electrical crossbar, allowing reduction of control network average hop count to a single hop (all node are connected via the crossbar). It optical network uses a ring topology more adapted for photonic interconnects, reducing the losses of waveguide crosses.

And finally a fully optical ring network-on-chip (FORNoC) is proposed. The architecture has the advantage of removing the need of higher power consumption electrical control network. The proposed FORNoC takes advantage of wavelength division multiplexing (WDM) by combining static and dynamic wavelength alloca-

tion techniques in the same architecture. A different wavelength-channel is statically allocated to each destination node for light weight communication. For heavy load communication, a shared multiwavelength-channel is available by requesting it in execution time from source node to a special node that manages dynamic wavelength allocation. Using cycle accurate and optical network simulator, we evaluate our proposals in terms of energy consumption and performance using probabilistic traffic patterns.

1.3 Thesis Overview

The rest of this thesis is organized as follows :

- In Chapter 2 an overview of on-Chip optical interconnects and related works are presented.
- Chapter 3 describes the proposed low latency path setup control network for a hybrid photonic torus network.
- In Chapter 4, we introduce a new hybrid electronic-optical network-on-chip consisting of an electrical crossbar path setup network and an optical ring data transfer.
- Chapter 5 presents a fully optical ring network-on-chip that uses static and dynamic wavelength allocations.
- In Chapter 6 we evaluate and discuss the performance of the proposed fully optical ring network-on-chip.
- Finally chapter 7 presents the conclusion of this thesis and outlines future research directions.

CHAPTER 2

OVERVIEW OF ON-CHIP OPTICAL INTERCONNECT AND RELATED WORKS

On-chip optical interconnect is a new field and remains less understood than electrical interconnect. With recent nanophotonics technology remarkable advances, optical interconnect is becoming a good candidate to replace their electrical counterpart to face the challenges of future CMPs and SoCs interconnection infrastructures. In this chapter we present an overview of on-chip optical interconnect and introduce some related works.

2.1 Optical Devices

Photonic network has been widely accepted as alternative to electrical one because it can be much more faster and energy efficient. In addition, optical link (waveg-

uide) with wavelength division multiplexing(WDM) allows the transfer of many information on multiple wavelength simultaneously which can increase the inter-connection bandwidth significantly. To utilize optical interconnect onto chip architecture however, only limited choice of materials and processes are available for fabricating optical components [8]. Also optical interconnects lack buffering and processing optical data in-flight features that limit their fully integration onto the chip level to replace completely their electrical counterpart. Figure 2.1 shows a block diagram of optical interconnect for on-chip usage.

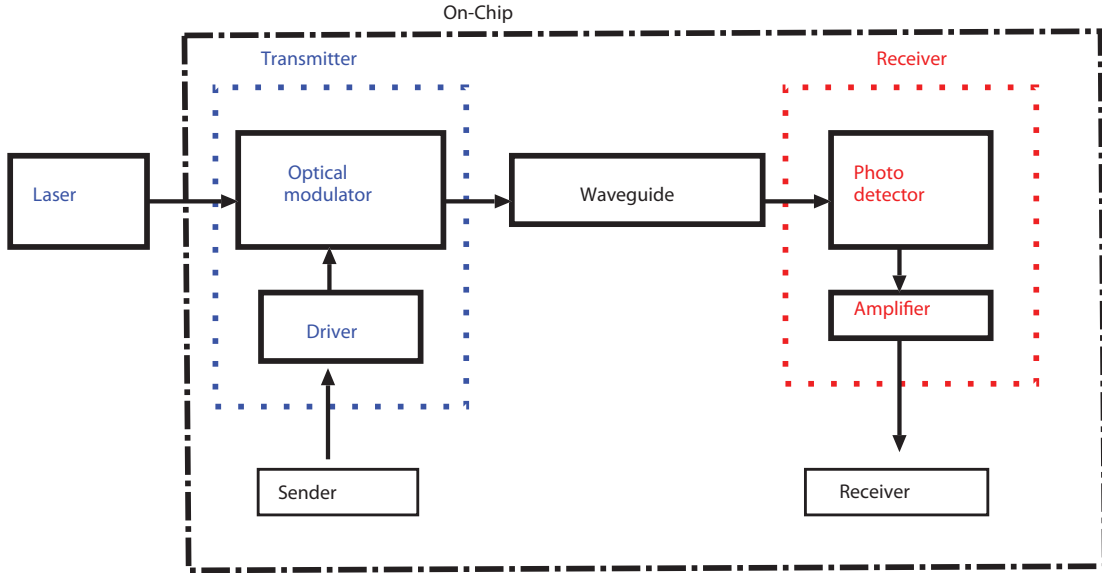


Figure 2.1: Block diagram of on-chip optical interconnect.

2.1.1 Light Source

In the Figure 2.1, the light source (laser) is supposed to be off-chip due to the absence of efficient silicon-based laser that can be monolithically integrated inside the chip. The light is then coupled into an on-chip waveguide that distributes light over the entire die. Using an optical modulator, the light is converted onto data (light

pulse) generated by a driver at the source node. The optical signal is then routed to a waveguide. At the receiver side, a photo detector convert the light pulses into a photocurrent. The photocurrent is then transformed into a conventional digital voltage signal by a trans impedance amplifier . Many such interconnects could be fabricated on the chip, their number being limited by available optical power, waveguide spacing limitations, detector and modulator area, as well as routing constraints [11].

2.1.2 Waveguide

The waveguide is a basic silicon photonic device which is used for carrying high-speed optical data from one node to another. Comparing to electrical links, optical waveguides have intrinsic advantage of high speed of light at lower energy cost [30]. Silicon photonic waveguides are able to transfer multiple wavelengths of optical data stream simultaneously. Furthermore, photonic waveguides can be bended, crossed, and coupled [15] from one to another in order to improve the flexibility for optical data transfer.

Recently, crystalline silicon waveguides with submicron dimension are considered as potential choice but has obvious insertion losses caused by physical crossings. Unlike crystalline silicon, deposited silicon nitride waveguide is placed as carrier medium in high speed communication links with the vision of monolithic integration of high performance. It has low crossing insertion losses and enormous potential for photonic links [1].

2.1.3 Modulator

The modulator is an essential component that used for high speed of conversion from electrical data to optical data. The laser source provides light source for the modulation. According to the electrical command data, the modulator is switched

“ON” or “OFF” to generated a sequential optical data in the waveguide by using light source(Electrical/Optical Conversion). The speed of modulation up to 12.5Gbps has recently been proved [21]. By using wavelength division multiplexing(WDM) technology, it is preferable to have wavelength-selective modulators that can encode data on multiple wavelengths and form a cohesive parallel optical data stream within a single waveguide. WDM technology helps the modulators achieve high bandwidth modulation for photonic NoCs.

2.1.4 Detector

The detector is placed at the destination of optical communication link for converting incoming optical signal into electrical domain(Optical/Electrical Conversion). Selective detectors, consisted of CMOS-compatible Germanium(Ge) doped resonant rings [20], can be used for receiving and translating different specific wavelengths. Ge-doped detector have demonstrated speed of detection up to 40Gbps.

2.2 Related Works

As nanophotonics has several advantages for on-chip applications, there exists considerable previous works on nanophotonic research that have shown several network designs that can overcome the limited bandwidth and high power dissipation of electrical interconnects. A few of these networks will be explained in more detail: Hybrid Circuit-Switch [25], CORONA [28], Firefly [20], FlexiShare [19], PROPEL [18] and all optical routed wavelength architecture [27] . Tabble 2.1 summarizes some of the properties of these architectures.

2.2.1 Hybrid Circuit-Switch

The Circuit-Switch network uses a simple electrical network to setup and tear down a high speed optical circuit switched torus network. In the Circuit-Switch network, when a source tile needs to communicate with a destination tile, the source tile will send electrical data through a circuit setup network that activates the correct micro-ring resonators for guiding the optical data to the correct destination. After the destination tile receives the optical data, the destination tile sends electrical data in reverse (from destination to source) that tears down the optical network. The issue with this network is the increased delay for setting up the optical circuit network for small data packets and increased blocking delay due to contention for shared channels. To reduce path setup latency of such architecture we propose a predictive switching and reservation based path setup network to improve its performance. Details about the architecture will be described in Chapter 3.

2.2.2 CORONA

The CORONA network is a 256-core optical bus network. CORONA's optical bus is constructed by using 64 multiple write single read (MWSR) nanophotonic channels, where many tiles can write onto the optical channel but only one tile can read the channel. In order to prevent two or more tiles from communicating at the same time, CORONA uses optical tokens to only allow one tile to communicate at a time. An optical token is a burst of optical light that circulates through all the communicating tiles. When a tile needs to communicate with the destination tile, it will activate a micro-ring resonator and try to capture the circulating optical token. The issue with CORONA is the high contention for optical tokens when two or more tiles require to communicate to the same destination. At the difference of CORONA, we propose an optical ring network with a dedicated statically allocated single-wavelength path to each source node, and a dynamic allocation multiwave-

length path shared among the nodes. Our architecture described in Chapter 5, combines both wavelength allocation techniques to reduce contention in the network resources.

2.2.3 Firefly

The Firefly network is an electro-optical network, that uses cheaper electronics to route data to local tiles, and nanophotonics to route data to global tiles. Optical channels within Firefly are constructed using single write multiple read (SWMR) nanophotonic channels. In a SWMR nanophotonic channel, a single tile can write on the optical channel but multiple tiles can read the channel. To prevent tiles within a SWMR nanophotonic channel from receiving a signal that is not destined for them, Firefly implements a reservation system that activates micro-ring resonators to guide the optical signals to the correct destination tile. Firefly strikes a balance between cheaper electronics for local communication and nanophotonics for global communication. However, the issue with Firefly is the higher power dissipation required for data to traverse over the electrical network and the latency penalty due to the reservation system.

2.2.4 FlexiShare

The FlexiShare network [19] is an optical crossbar network that combines the benefits of MWSR nanophotonic channels with SWMR nanophotonic channels. MWSR nanophotonic channels allow for multiple tiles to write data on a single communication channel but only one tile can read the data. SWMR nanophotonic channels allow one tile to write data and several tiles to read the data at once. By combining the benefits of MWSR and SWMR nanophotonic channels, a tile can use any nanophotonic channel to transmit data to any tile. To prevent two or more tiles from transmitting on the same nanophotonic channels, FlexiShare uses optical to-

kens similar to CORONA. Once a tile captures an optical token, FlexiShare uses a technique similar to Firefly's reservation system to prevent the incorrect tiles from receiving the data. The major advantage of FlexiShare is the ability to reduce the number of nanophotonic channels used in the network, as each nanophotonic channel is connected to all the tiles. The issue with Flexishare is the high number of micro-rings resonators required for each nanophotonic channel and the high optical losses due to long waveguides. FlexiShare also requires contention resolution from both sender and receiver side. Multiple sender/receiver can use the same receiving/sending optical channel.

2.2.5 PROPEL

PROPEL is a 64 core NoC that strikes a balance between electronic and photonic interconnects [18]. Nanophotonic interconnects are used for long distance inter-router communications, while electronic switching and flow control are used for nodes within the same tile. In addition of using different topology from FIREFLY, PROPEL statically allocates optical channels for long interrouter communications. As in Firefly, the issue with PROPEL is the higher power dissipation required for data to traverse over the electrical network.

2.2.6 All Optical Routed Wavelength Architecture

This architecture uses a passive routing of optical data streams based on their wavelength, the architecture eliminates the need for optical resource reservation [27]. Unfortunately, bandwidth performance is limited due to the allocation of wavelengths to specific source-destination pairs. The issue with this architecture is that all path are source-destination specific. Although the architecture almost eliminate the need for contention resolution, many network resources could be idle most of the time.

Table 2.1: Related works

Related Works	Network Topology	Wavelength Allocation	Interconnect Type
Hybrid Circuit-Switch	2-D Torus, Mesh	Dynamic	Electrical path setup(control), and optical data transfer
CORONA	Crossbar	Static	Fully optical
Firefly	Butterfly	Static	Electrical (within clusters) and optical for extra links
FlexiShare	Crossbar	Dynamic	Fully Optical
PROPEL	Modified mesh with extra links	Static	Electrical for the mesh network, and optical for the extra links
All Optical Routed Wavelength	2-D-HERT(Hierarchichal Ring Topology)	Static	Fully optical

CHAPTER 3

LOW LATENCY PATH SETUP HYBRID PHOTONIC TORUS

In this Chapter, we present a low path setup hybrid torus NoC using predictive switching [16] and a reservation based path setup techniques for the electrical control network to reduce the setup latency of a conventional hybrid photonic torus. Since the circuit setup latency plays a key role in the overall performance of HPNoC [26], we use these techniques to reduce the path setup latency thus improving the overall network performance.

3.1 Hybrid Photonic Torus

While the hybrid photonic NoC offers unique advantages in terms of bandwidth and energy compared to fully electrical NoC, its implementation requires extra hardware to support the optical communication such as : light source (laser), modulators, waveguides, optical switches, and demodulators [26]. Fig. 3.1 shows a 4×4

Chapter3. Low latency Path Setup Hybrid Photonic Torus

torus HPNoC. The topology consists of 2 layers: an optical high-bandwidth data transfer circuit switching network, and an electrical packet switching control network. Nodes in the HPNoC communicate as follows:

- Firstly a path setup message is sent by the source node in the electrical network to establish a path for the optical network.
- After the path is set, an acknowledgment pulse is sent back to the source node by the destination node in the optical network, and optical data can be transferred without need for buffering at intermediate nodes.
- Finally when all data are sent, a teardown message is sent by the source node in the electrical control network to release the optical circuit.

Similarly to a circuit switching flow control, the HPNoC performs better with larger message sizes because of the high speed data transfer in the optical network once the communication path is established. When only a few small-sized data transmissions occur, the HPNoC is not needed, while a cheap simple electrical NoC fits with such a case.

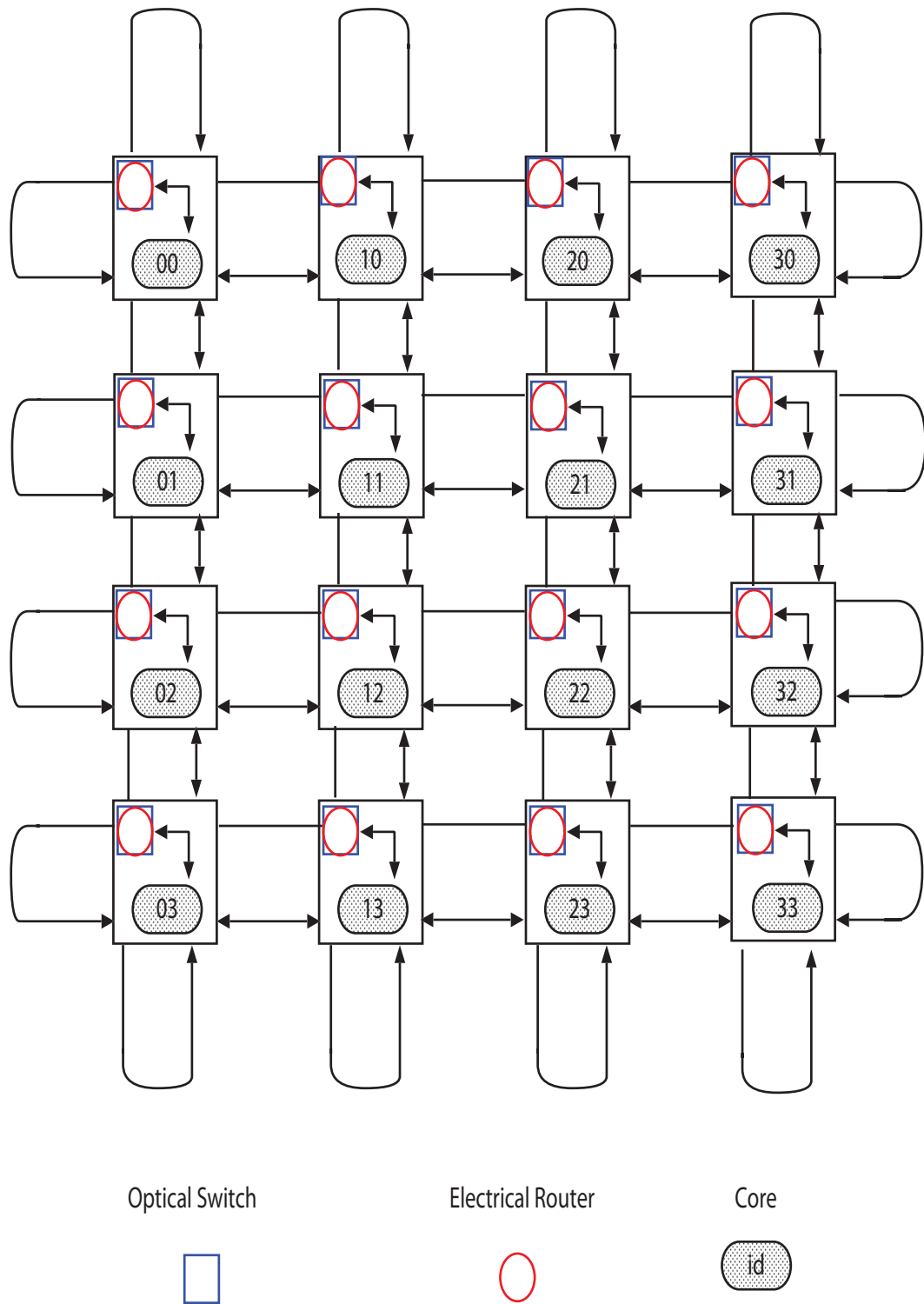


Figure 3.1: A 4x4 hybrid torus photonic NoC

3.1.1 Optical Network

The optical network comprises optical switches connected by optical waveguides. At each node, an optical modulator and detector are needed for electrical-optical-electrical (E-O-E) conversions. At the source node, an external laser light is modulated in the optical modulator from electrical to optical data signal. The modulated optical signal is transmitted on the optical waveguides. At the destination node, the optical signal is detected by the optical detector and ejected from the optical network. To build a 2D torus topology, a 5×5 optical switch is necessary for each node: one input/output port for each direction (WEST, NORTH, EAST, and SOUTH) and one for the processing element. To remove the need for extra injection and ejection gateways in the switch used in [24], we use the optical switch proposed in [7] shown in Fig. 3.2. The switch consists of micro-ring resonators, waveguides and a control unit. By turning ON/OFF the state of a resonator, light can be directed in the switch from one direction to another according to the control unit which is set by the electrical network. For instance in Fig. 3.2(a), optical data coming from the GATEWAY port is guided to the WEST output port by turning “ON” the resonator 4. The same data can be guided to the EST port by turning “ON” the resonator 2 shown in Fig. 3.2(b).

The high bandwidth capabilities of optical interconnects are due to the use of WDM. It statically allows the transfer of optical data using all wavelengths within a waveguide for the same source-destination pair’s data stream. Optical switch with a smaller number of micro-ring resonator presents a better solution for hardware cost. The optical switch we used only required 12 micro-ring resonators. To implement a dynamic allocation(wavelengths of the same waveguide is divided among multiple data stream), however, the cost of the optical switch increases. The number of resonators is multiplied with the corresponding number of wavelengths used as each micro-ring resonator uses a unique resonance wavelength. The arrangement

of the waveguides and micro-ring resonators made this optical switch suitable for mesh and torus networks that use dimension order routing (DOR). It removes unnecessary turns that are avoided in DOR.

3.1.2 Electrical Network

The electrical control network consists of electrical routers interconnected by electrical wires in a torus topology. We propose two path setup techniques to improve the performance of the control network by reducing the electrical network latency.

3.1.2.1 Predictive Switching Based Path Setup

For the predictive switching based path setup, we use prediction routers. The hardware area of the electrical network is increased by 4.8-12.0% as reported in our previous work [17]. Prediction routers speculatively forward the packets inside a router bypassing some pipeline stages. The prediction router is shown in Fig. 3.3(b). The differences from the conventional router shown in Fig. 3.3(a) are as follows:

- 1) A predictor is added at each input channel.
- 2) The arbitration unit for virtual-channel and switch allocations (VSA Arbiter) is modified to handle the tentative reservation from predictors.
- 3) And a kill signal is added at each output channel in order to remove miss-routed flits when the prediction fails [16].

The predictor in an input-channel forecasts which output channel will be used by the next packet transfer before it reaches the input-channel. Then it asserts the reserve signal to the arbiter in order to tentatively reserve a time-slot of the crossbar for the predicted output-channel. The VSA arbiter handles the request and reserve signals from each input-channel(configure). If the prediction fails, the kill signal is asserted to the miss-predicted output channel. The output-channel will mask all incoming data as dead flits (miss-routed flits) which never propagate to the outside of the router. With this technique, when the prediction hits, it is possible to complete

the switch traversal (ST) within one router cycle and bypass the pipeline stages of routing computation (RC), virtual-channel allocation (VA), and switch allocation (SA) which are required in the conventional router [6]. When the prediction fails, the conventional packet processing is carried out. It is important to note here that there is no miss-penalty on the miss-routed latency.

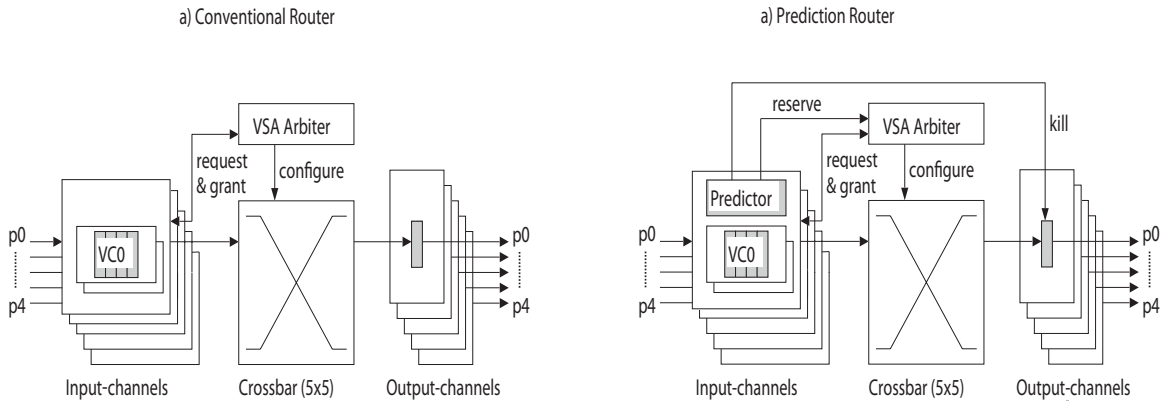
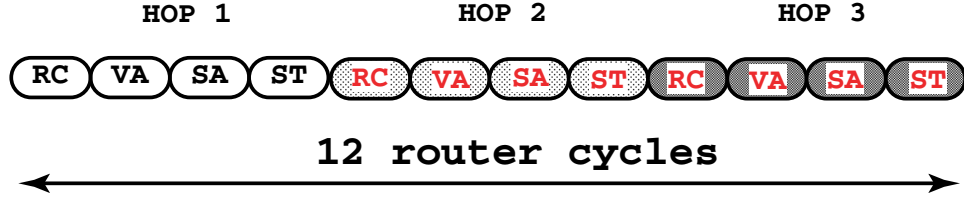


Figure 3.3: Electrical routers

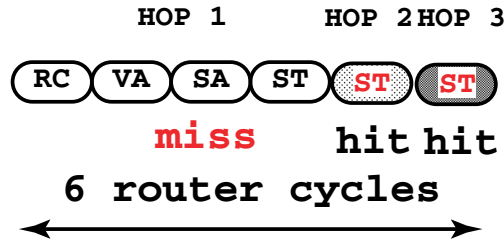
Figure 3.4 as an example compares a timing diagram for sending a packet through 3 hops using a conventional router (Fig. 3.4(a)) and the prediction router for the electrical control network (Fig. 3.4(b)). With the prediction router, the *end-to-end-latency* is reduced by half from 12 router cycles, necessary in the conventional router, to only 6 cycles in the case of the predictions hit in two of the 3 hops.

By processing packets before they arrive at input buffers using look-ahead routing, only a single stage pipeline (ST) is necessary for packet transfer when prediction hits. The prediction mechanism, therefore, drastically reduces the packet processing latency per router. If a switching with high prediction hit rate is applied to the electrical control network of the HPNoC, it is possible to decrease the circuit setup latency and improves its overall performance.

Since some pipeline stages are skipped only when the prediction hits, the pri-



(a) Normal Router



(b) Prediction Router

Figure 3.4: Pipeline time diagram for conventional and prediction routers

major concern for reducing the communication latency is the prediction accuracy. We use the following two prediction algorithms.

- *Latest port matching (LP)*: The LP strategy predicts in such a way that the next incoming packet will be forwarded to the same output-channel as that of the previous packet. The LP predictor requires only a single history record in each input-channel, leading to a lower hardware overhead cost.
- *Sampled pattern matching (SPM)*: The SPM algorithm was originally proposed as a universal predictor [10]. It selects a value with the highest probability after a suffix sequence, called a marker, in a given data set. The predicted value is calculated by using the majority rule to all values appearing at positions just after the markers in the data. We can use it to predict an output-channel for the next incoming message of an input-channel by finding the most fre-

quently used output-channel after the longest suffix sequence (marker) of the communication history. An example of prediction using the SPM prediction mechanism is shown in Fig. 3.5. In step 1 of the algorithm, the marker is determined by finding the longest repeated sequence from the history of past used output-channels used by an input-channel, in this example the marker is “0012”. Second, the values appearing at positions just after the markers in the history are recorded and counted (Step 2). Finally in Step 3, the predicted value is calculated by applying a majority rule to all values of Step 2. Here, since value “3” appears one time and value “2” appears two times, the predicted value is “2”.

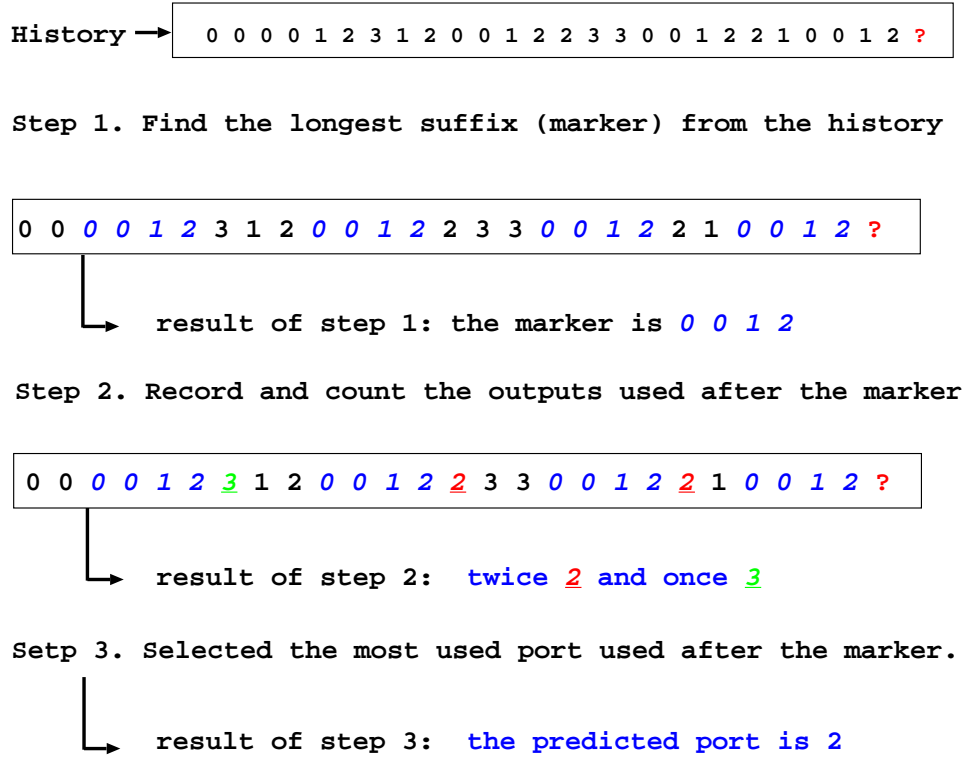


Figure 3.5: Example of prediction using SPM scheme

3.1.2.2 Reservation Based Path Setup

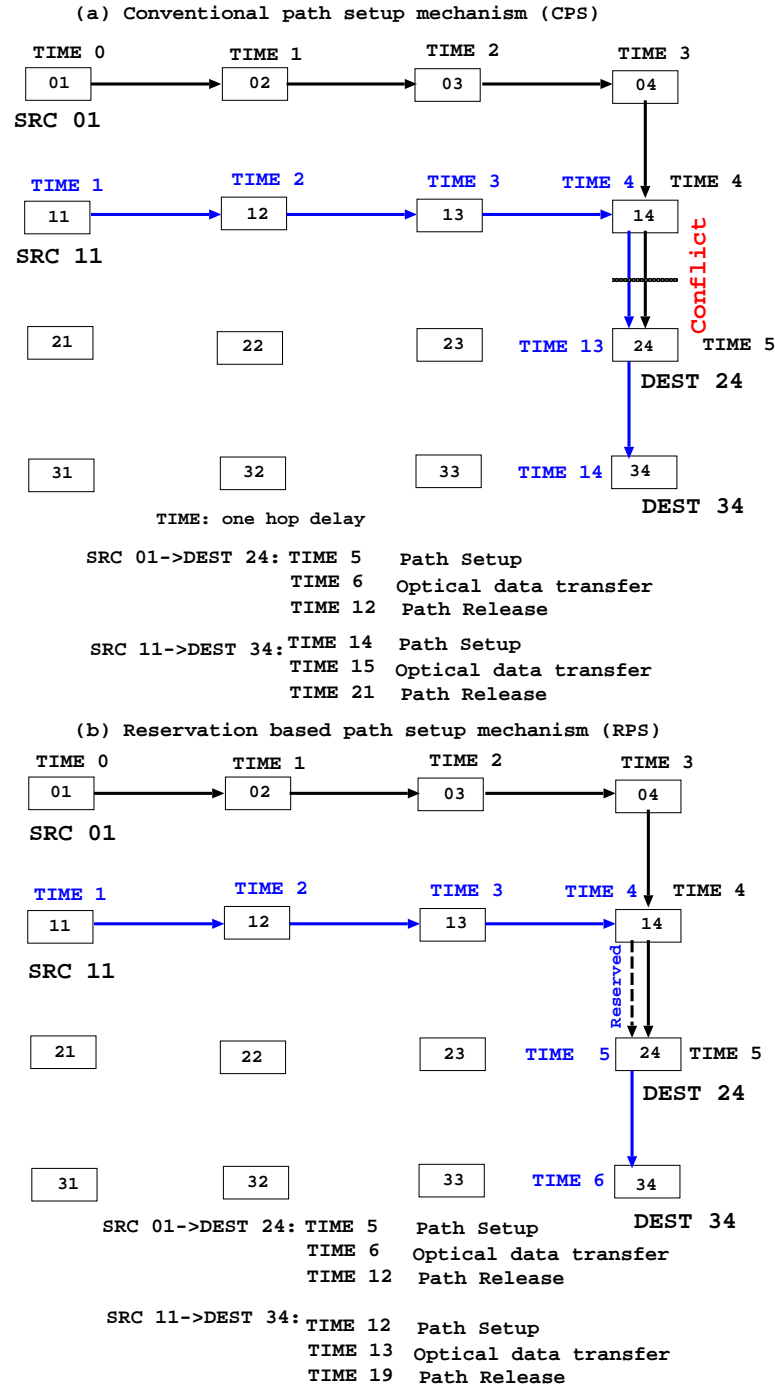


Figure 3.6: (a) Conventional vs (b) reservation based path setup mechanisms

A contention resolution mechanism is required when several path setup messages compete for the same path or a portion of a path. It directly affects the performance of the setup latency. For the prediction technique, we implement the simplest contention resolution mechanism, we called conventional path setup (CPS) shown in Fig. 3.6(a). In this case when two path setup messages for the same portion of a path (path between node 14 and node 24), one of them is granted the path (communication between node 01 and node 24) and the other one is buffered until the path becomes available. The source-destination pair (11, 34) will set the path after its release by pair (01,24). The two source-destination communications finish at TIME 21.

In Fig. 3.6(b) we propose a reservation based path setup (RPS) mechanism. In this technique, the ungranted path setup message of the source-destination pair (11,34) instead of being buffered at node 14 where there is a path-conflict, it reserves the path and moves toward the destination. The release path message of the pair (01,24) sets the reserved path for communication at TIME 12. The two pair communications finish at TIME 19. Their latency for communication is reduced by two hop latencies. As shown in this example, the reservation mechanism also can reduce the path setup latency and improves the end-to-end communication latency in the HPNoC . To implement RPS, the electrical arbiter hardware of the conventional electrical router is slightly modified for handling path reservations. RPS only reduces path setup latency when contention occurs in the communication patterns. For traffic patterns such as neighbor in which node trends to communicate with their adjacent nodes, both CPS and RPS performs similarly. Table 3.1 summarizes the advantages and disadvantages of both path setup mechanisms.

Table 3.1: Advantages and disadvantages of CPS and RPS

Path setup	Advantages	Disadvantages
CPS	-Simple arbitration scheme.	-Path setup messages are buffered when path conflicts occur.
RPS	-Reduction of latency when path conflict occurs.	-Extra arbitration required for handling reservation of paths.

3.2 Performance Evaluation

In this section First we compare the power consumption of an all-electrical NoC and a HPNoC, then we estimate the performance of our proposed path setup techniques for HPNoC.

3.2.1 Power Consumption Estimation

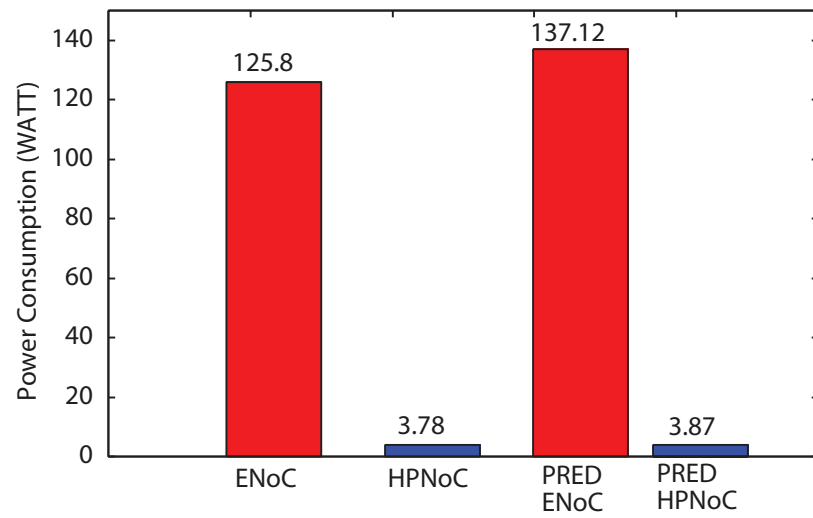


Figure 3.7: Power consumption cost

Chapter3. Low latency Path Setup Hybrid Photonic Torus

The main motivation of using photonic NoC is its potential to reduce the high power consumption of an electrical NoC to provide the same performance for intra-chip communications. To offer the same performance of a photonic NoC, electrical NoC requires the use of many parallel links leading to a higher power dissipation of the network.

By scaling the power cost calculation method used in [23] to our 64 nodes torus network we evaluate the power consumption of the electrical and HPNoC.

In the Electrical NoC, the total energy consumed by the network can be computed as :

$$E_{NETWORK-CYCLE} = \left(\sum_{j=1}^{N_L} U_{Lj} \times E_{FLIT-HOP} \right) \times f \quad (3.1)$$

where U_{Lj} is the average number of flits traversing link j per clock cycle, an estimate on the utilization of link j ; $E_{FLIT-HOP}$ is the sum of energy spent by a flit in the different pipeline stages of flits processing; and f the clock frequency of the router.

For the HPNoC, the dissipated energy is estimated as the sum of the energy of two components: the photonic network, and the electrical control network.

- Since the electrical control network differs from the conventional electrical NoCs in terms of message size, the energy can be deduced from the electrical NoC's one using the same equation (3.1) scaled to the electrical control message size.
- The energy consumed by the photonic network consist of :
 - 1) The transmission energy which is calculated as :

$$P_{P-NoC,transmission} = NR_{ON-STATE} \times 0.5mW \quad (3.2)$$

Where $NR_{ON-STATE}$ is the number of micro-ring resonators in "ON" state, and $0.5mW$ is the assumed energy cost for a single micro-ring resonator in

“ON” state [23]. No energy is consumed by an “OFF” state micro-ring resonator.

2) And the modulator/demodulator energy is estimated as:

$$P_{P-NoC,mod/demod} = 0.11pJ/bit.64.Bandwidth \quad (3.3)$$

We compute the energy consumed by a HPNoC and a fully electrical NoC for a 32 nm node technology that uses a 5 GHz router clock frequency to provide the same performance bandwidth. By assuming an average link utilization of 50% for the 64 nodes torus of 800Gbps data transmission bandwidth, we estimated the energy consumed by the two networks. When using the prediction router the energy consumed is majorated by an extra 9% of the electrical network energy due to the extra overhead added by the prediction router [16].

Fig. 3.7 plots the power estimation results. It shows that the electrical NoC consumes a huge amount of power compared to the HPNoC to be able to deliver the same bandwidth performance. It further shows that the extra energy overhead required when using the prediction router is almost neglectable for the HPNoC.

3.2.2 Simulation Conditions

We evaluate the performance of the networks using a modified version of the booksim [6] cycle accurate simulator. For simulation, we use three probabilistic traffic patterns :

- *Uniform random* : Each node sends a packet to a randomly chosen node.
- *Neighbor* : Each node sends a packet to its neighboring nodes.
- *Bitreversal* : Each node sends a packet to a destination whose address is the bitreversal of the sending node address.

The Table 3.2, 3.3, and 3.4 summarize our simulation parameters.

Table 3.2: Simulation parameters

Simulated Networks	<i>ENoC (w/wo prediction), HPNoC (w/wo prediction), HPNoC (CPS, RPS)</i>
Topology	<i>2D Torus 64 nodes</i>
Routing	<i>DOR</i>
Control message size	<i>4 Bytes</i>
Data size	<i>20 Bytes</i>
Prediction algorithms	<i>LP, SPM</i>
Traffic patterns	<i>Uniform, Neighbor, Bitreversal</i>

Table 3.3: Optical NoC parameters

Number of wavelengths per waveguide	64
Data rate per wavelength	12.5Gbps
Total link bandwidth	800 Gbps

Table 3.4: Electrical NoC parameters

Router frequency	5 GHz
Number of VC per physical channel	2
Channel width	32 bits
Buffer size/VC/channel	20 Bytes
Latency/hop without using prediction	4 router cycles
Latency/hop when prediction is used and hit	1 router cycle
Latency/hop when prediction is used and miss	4 router cycles

3.2.3 Results and Discussion

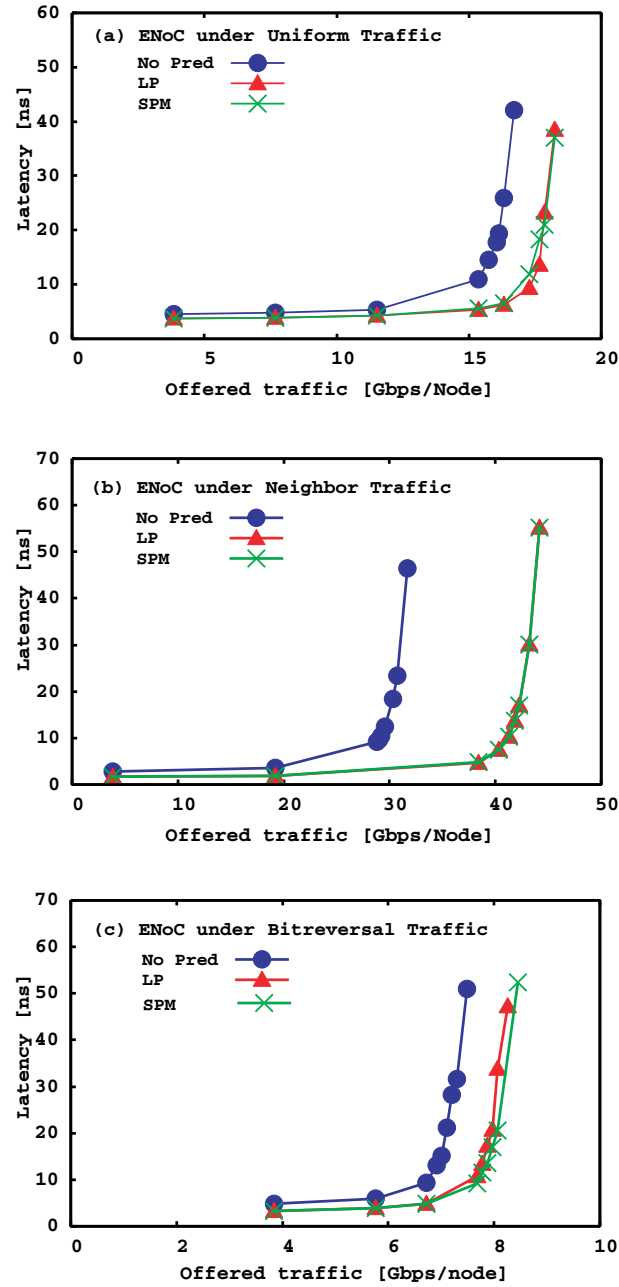


Figure 3.8: Electrical NoC under uniform (a), neighbor (b), and bitreversal traffic patterns (c)

The predictive switching and RPS are techniques to reduce the latency path setup messages spend in the electrical control network. By reducing the average path setup latency, the control network with this techniques can afford more messages before network saturation thus improving the overall performance of the HPNoC.

Fig. 3.8 (a), (b), and (c) show the simulation results for a fully electrical network under uniform, neighbor, and bitreversal traffic patterns, respectively. The results show that both LP and SPM prediction techniques improve the performance of the network for all traffic patterns. For instance, using the prediction router, the electrical NoC can be loaded with nearly an extra 10 Gbps/node compared to the conventional electrical one for the neighbor traffic pattern as shown in Fig. 3.8 (b). In the case of uniform traffic patterns, due to the random communication pattern, LP and SPM schemes show nearly the same performance as seen in Fig. 3.8 (a). In the case of neighbor traffic pattern, due to the fact that nodes trend to communicate with their adjacent nodes, the LP scheme obtains nearly the same prediction hit rate as the SPM, leading to almost the same improvement of latency as shown in Fig. 3.8 (b). As seen in Fig. 3.8 (c), SPM prediction technique shows better performance than LP under bitreversal traffic pattern due to the analysis on the longer output history used by an input channel of SPM.

In Fig. 3.9 (a), (b), and (c), the HPNoC performance is evaluated for uniform, neighbor, and bitreversal, respectively with and without LP and SPM prediction mechanisms. The results show that both prediction techniques improve the network performance. In particular for neighbor traffic pattern shown in Fig. 3.9 (b), this performance is almost doubled with the prediction techniques. Furthermore, these results also show that even with the simplest LP prediction technique which requires only a single output history at each input-channel, we can achieve a considerable increase in performance.

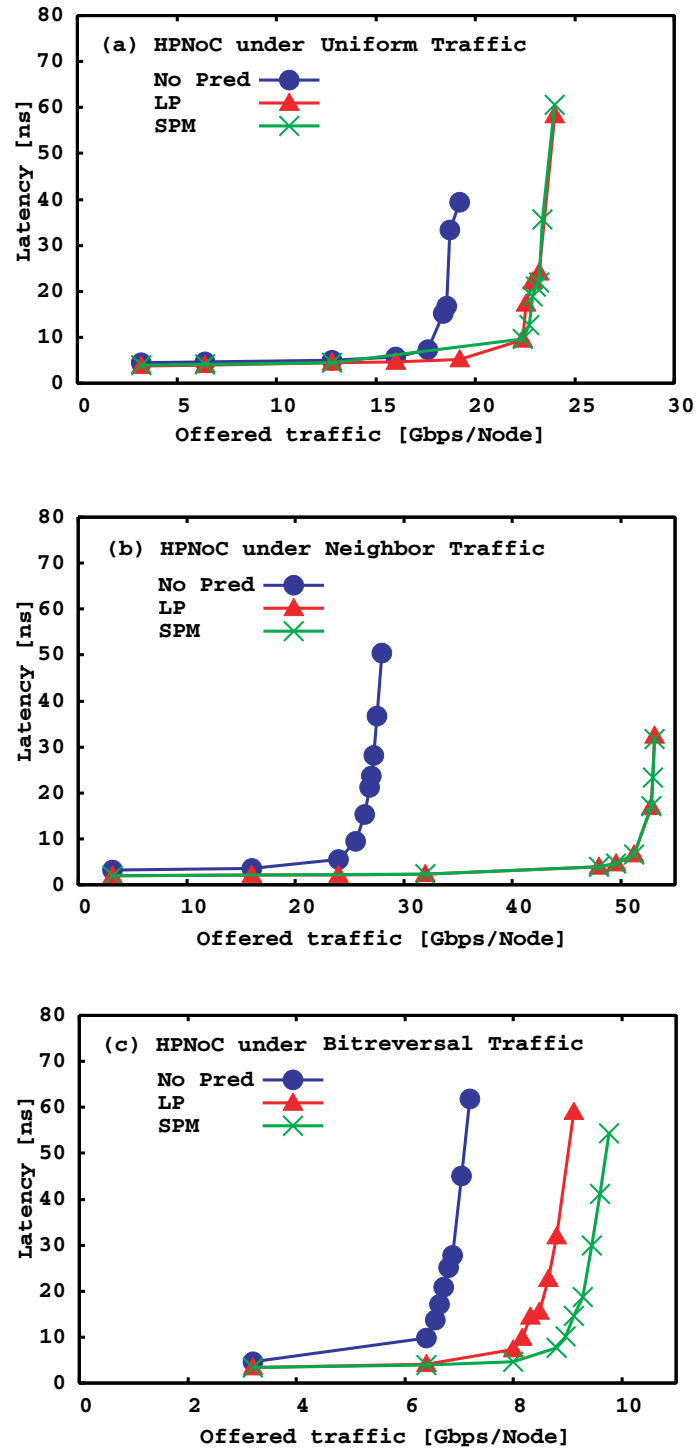


Figure 3.9: HPNoC under uniform (a), neighbor (b), and bitreversal traffic patterns (c)

Fig. 3.10 (a), (b), and (c) show a comparison of HPNoC against a fully electrical NoC under uniform, neighbor, and bitreversal traffic patterns, respectively with and without prediction technique. The results show that the HPNoC with the simplest LP predictive switching leads to better performance than all other simulated network configurations for all traffic patterns. Since the HPNoC uses a circuit switching flow control even for neighboring communication, a setup packet for establishing a path is necessary before communication can take place. The effect of path setup time for such communication pattern is particularly important in message delivery latency. That causes the packet switching ENoC without or with prediction outperforming the HPNoC without prediction as shown in Fig. 3.10 (b). However, by reducing the effect of path setup time the HPNoC with prediction outperforms all other configurations.

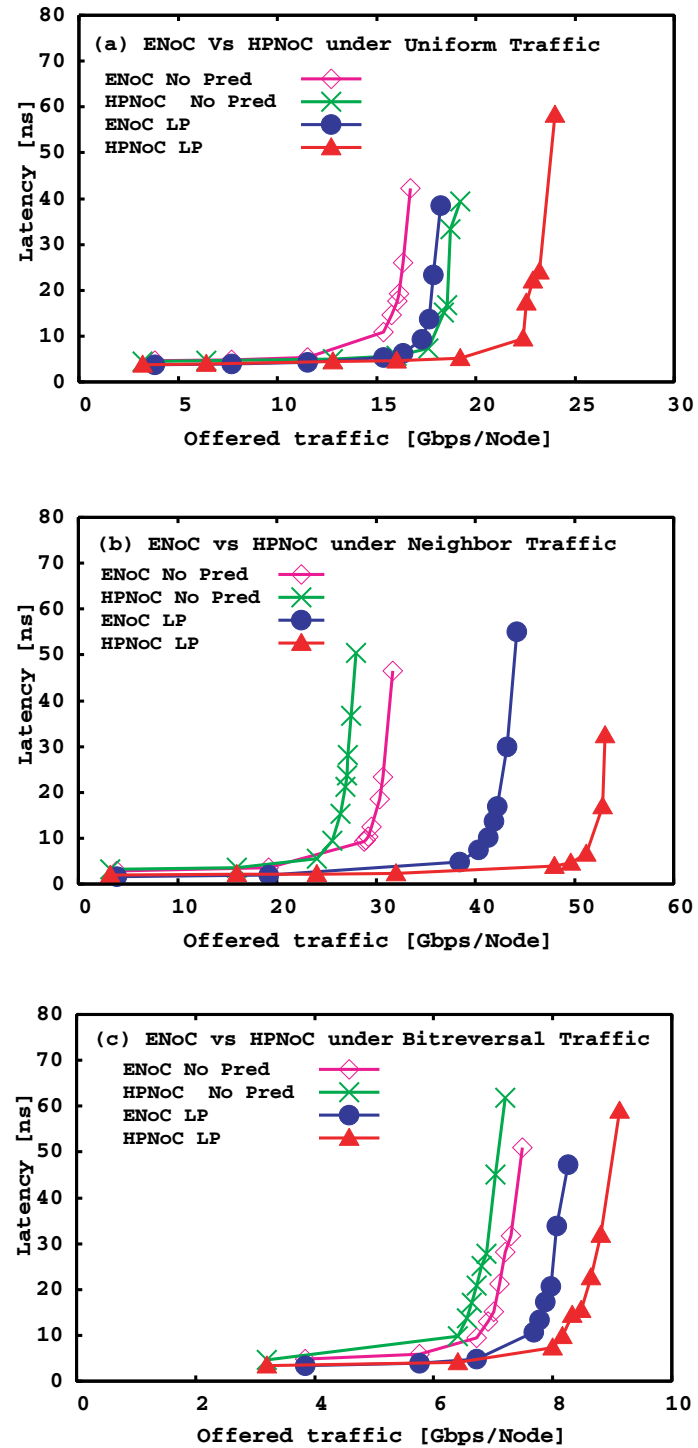


Figure 3.10: Electrical NoC vs HPNoC under uniform (a), neighbor (b), and bitreversal traffic patterns (c)

In Fig. 3.11 (a), and (b), we compare the performance of the conventional path setup (CPS) mechanism and our proposed scheme (RPS) for uniform, and bitreversal traffic patterns, respectively. Results show an improvement in all cases. By reserving the path ahead a time instead of buffering the path setup message, the average path setup latency is considerably improved leading to a better overall performance of the HPNoC.

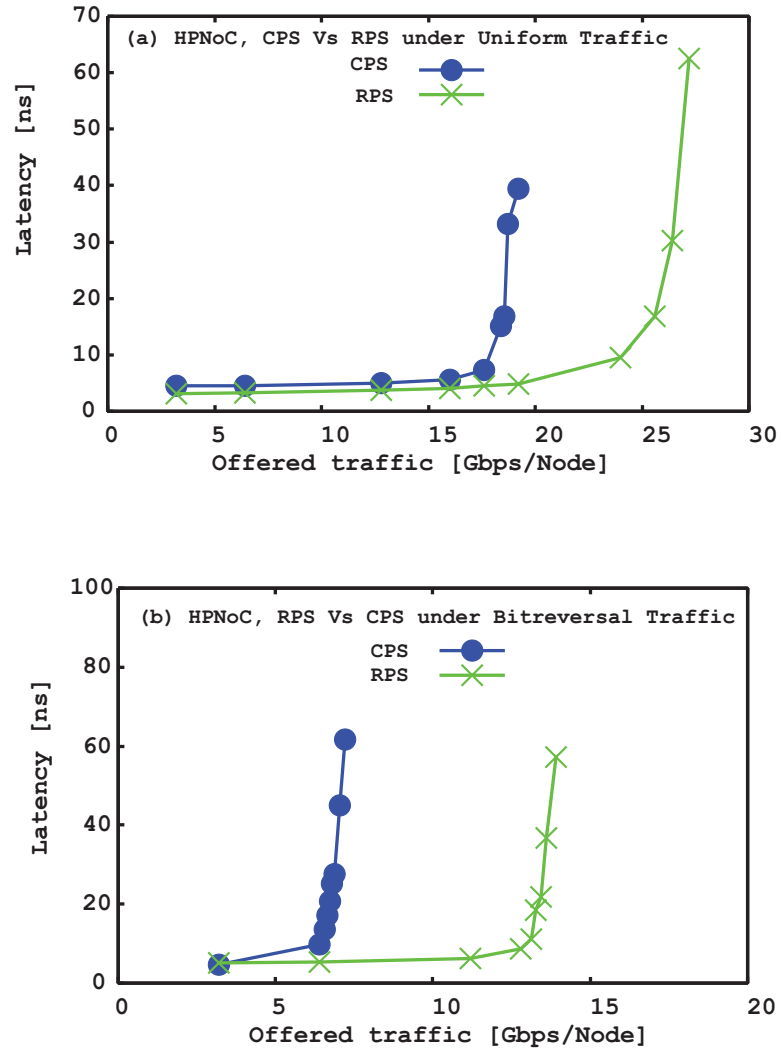


Figure 3.11: HPNoC, CPS vs RPS under uniform (a), and bitreversal traffic patterns (b)

3.3 Conclusion

Well designed optical interconnection has the potential to meet the high bandwidth and low power consumption required for future on-chip interconnection. In this Chapter, we have proposed path setup techniques to reduce the path setup latency for circuit switching HPNoC. The simulation results for probabilistic traffic patterns show that both techniques drastically improve the network performance of a conventional HPNoC. As crucial performance factor of the HPNoC is the setup time of the optical path, reducing the path setup latency in the electric NoC leads to a considerable gain in overall performance for HPNoC. In the next Chapter we further investigate an improved hybrid architecture which use different topology for the optical and electrical control networks.

CHAPTER 4

OREX: HYBRID OPTICAL RING ELECTRICAL CROSSBAR NETWORK-ON-CHIP

In this chapter we describe our proposed hybrid architecture consisting of an optical ring and an electrical crossbar central router (OREX). OREX takes advantage of both electrical and optical technology designs state-of-art to deliver a high data rate transfer NoC at an acceptable power consumption cost. An optical message is transmitted on the optical ring preceded by its path set-up performed by an electrical control packet switched using a crossbar switch. The crossbar switch is suitable to reduce the path set-up time compared to direct network topologies by reducing the path setup's hop count. Since the size of a control packet is small, we can restrict power consumption of the electric network. Latency as well as power consumption of the optical network is much lower than those of electrical one, so that total communication performance per power can be much higher than pure electrical net-

works. Another merit of the optical NoC is that wavelength-division multiplexing (WDM) enables simultaneous multiple messages transfer on a single waveguide. In order to use WDM, the crossbar switch host an allocator to perform wavelength allocation to nodes on the optical ring.

4.1 OREX Architecture

Figure 4.1 shows an 8 node OREX NoC's topology. OREX is a hybrid architecture of electrical and optical NoCs. It consists of an external laser source to provide light for modulating the data optically, nodes (small circles), an electrical central router for path setup, network interfaces (small rectangles), optical routers (big circles) electrical links, and optical links (waveguides).

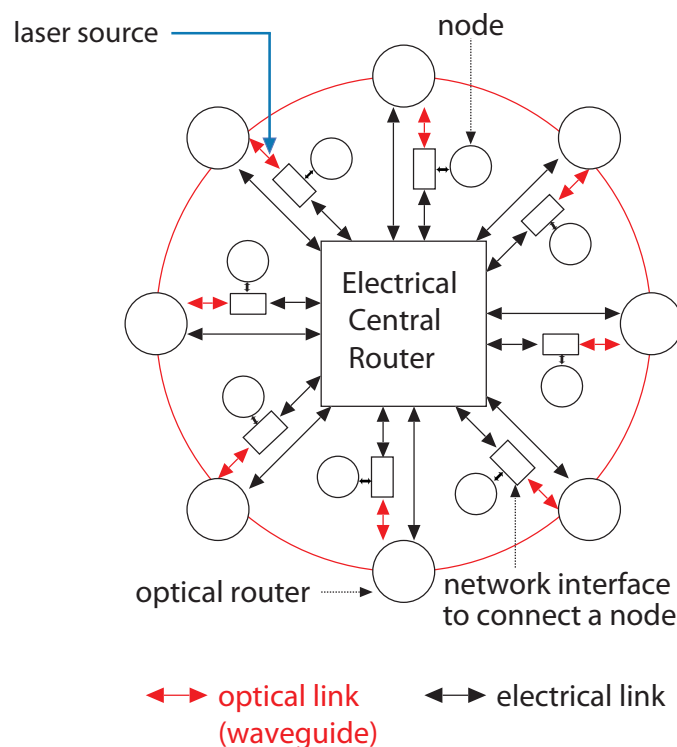


Figure 4.1: An 8 nodes OREX topology

Chapter4. OREX: Hybrid Optical Ring Electrical Crossbar Network-on-Chip

- *Network interface:* The network interface consists of a modulator for electrical-optical (EO) data conversion, and a light detector for optical-electrical (OE) data conversion. The external laser source provides the necessary light for data modulation. The network interface also serves as interface for connecting the node to the electrical central router for path setup.
- *Electrical central router:* The electrical central router consists of a $n \times n$ bidirectional input/outputs port crossbar switch where n is the number of node, an arbiter, and an optical path allocator. The optical path allocator is a unique unit which allocates optical paths, including wavelength assignment, between source-destination pair nodes.
- *Optical routers:* The optical router consists of optical switch formed by micro-resonators (MRs) which are placed at intersection of waveguides, and a controller of the MRs connected to the optical path allocator of the central router that sets the MR states. The MR has two states, “ON” and “OFF”. Depending on its resonance, a MR can be either dedicated to a waveguide, a group of wavelengths, or a specific wavelength. When the state of a MR is “OFF”, an input optical stream passes through the intersection, such as right to left, and vice versa. On the other hand, optical stream turns at the intersection when the MR is “ON” in order to receive/send optical data to/from destination and source nodes. At Initial state all MRs are “OFF” so that optical streams pass through on the ring at intermediate nodes. Therefore, we don’t need to change the state of the MRs at the intermediate optical routers. The electrical central router sends command to set “ON” the MRs at the source and destination nodes when an optical path is allocated to the communication pair. At the release process, they are reset to “OFF” state.
- *Optical link:* The basic OREX topology consists of two unidirectional waveguides forming a bidirectional link that connect the network nodes in a ring

topology. Each waveguide consists of multiple wavelengths. The optical link is divided into optical paths that may consist of a waveguide, a group of wavelengths, or a single wavelength.

Unlike a shared bus[12], OREX allows many simultaneous transfer along disjoint paths, such that the first node can send to the second node while the second node sends to the third, and so on. Figure 4.2 shows detailed connections between a node and optical/electrical routers for an OREX with two optical paths (clockwise and counter clockwise rings).

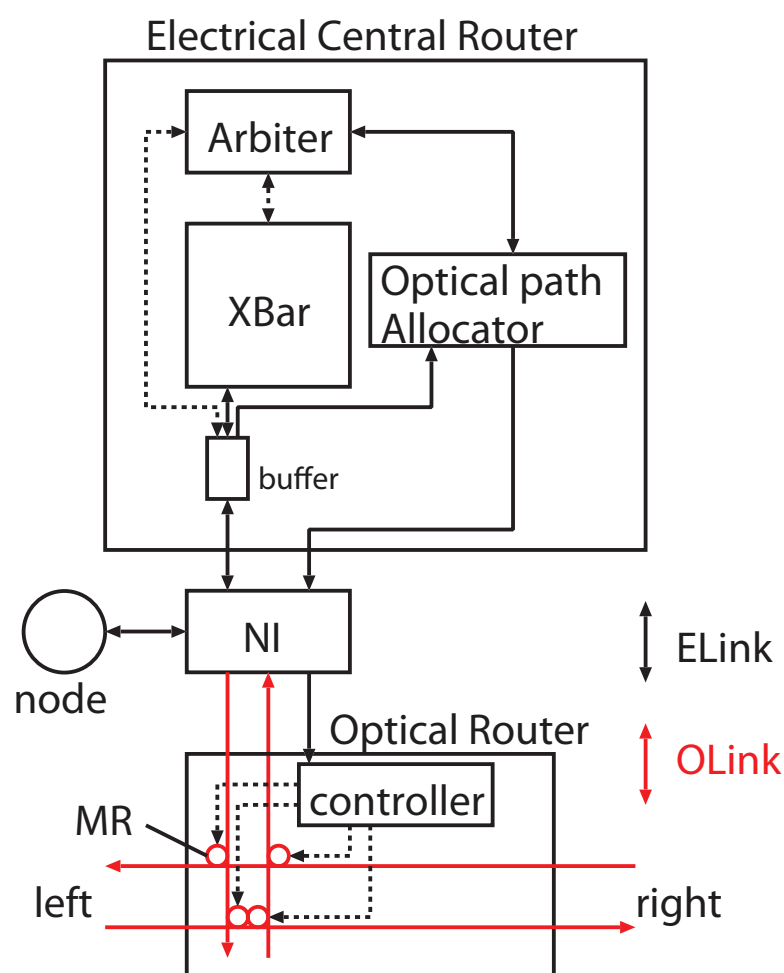


Figure 4.2: Connection between a node and routers

4.2 Communication Mechanism

OREX uses a circuit switching flow control. The communication mechanism consists of three steps: path setup, optical data transfer, and path release.

4.2.1 Path Setup and Optical Data Transfer

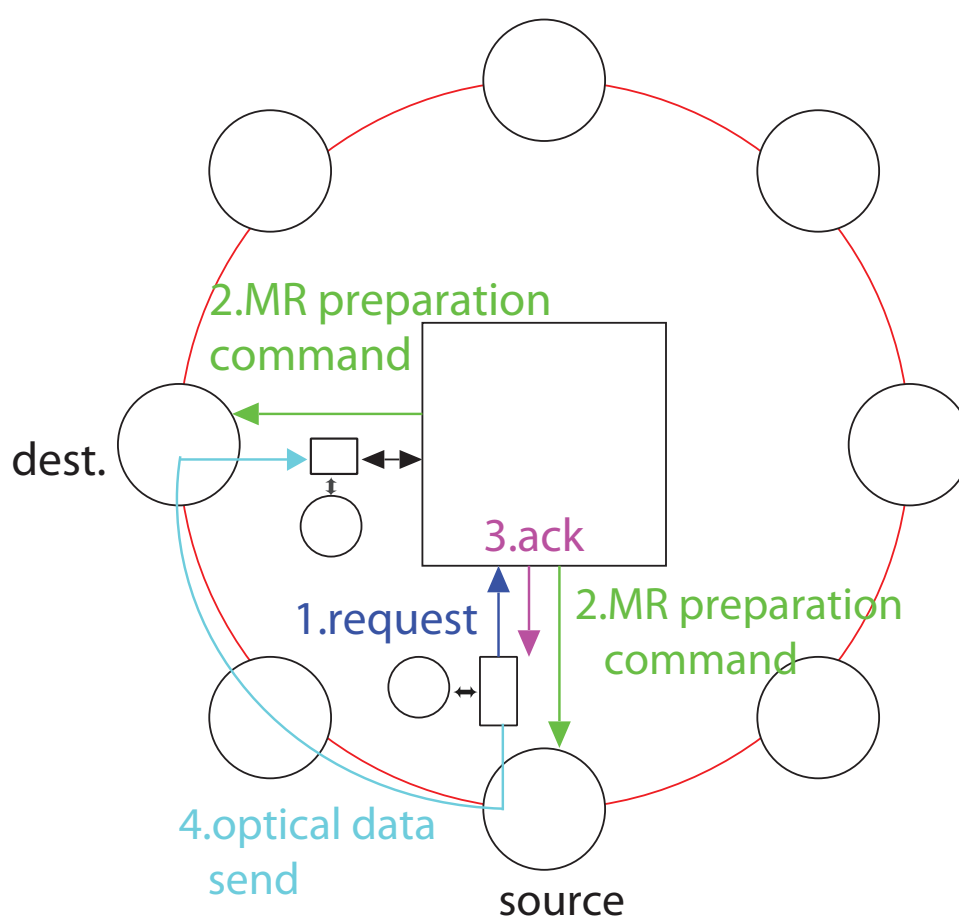


Figure 4.3: Path setup (1, 2, 3), and optical data transfer (4) sequences of OREX communication.

Chapter4. OREX: Hybrid Optical Ring Electrical Crossbar Network-on-Chip

Before any communication takes place, a full optical path¹ is reserved between the source and the destination. Figure 4.3 describes the path setup and data transfer mechanisms of OREX.

- The source node sends an electrical request packet to the central router via the network interface.
- The optical path allocator manages optical path assignment along the optical ring. When it successfully finds an available optical path between the source-destination node pair, commands are sent to both source and destination optical routers. Controllers inside the optical routers receive the commands, and set MR's state to "**ON**" in order to route the optical data.
- An acknowledge packet is returned to the network interface of the source node to notify the optical path establishment between source and destination.
- The source node then modulates the data to the optical path along the ring(optical data transfer).

¹portion of the ring (waveguide, wavelength, or group of wavelengths) used to connect source and destination nodes

4.2.2 Path Release

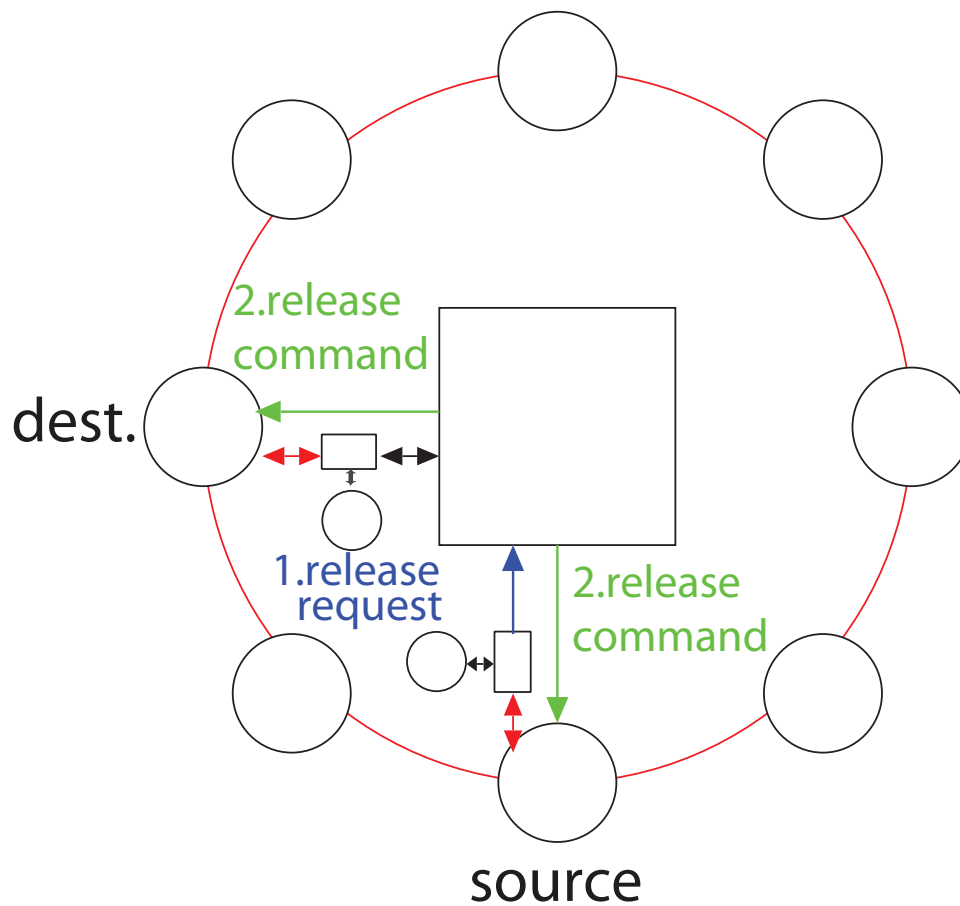


Figure 4.4: Optical path release processes

Figure 4.4 shows the path release processes of OREX.

- After transferring optical data, the source node sends release request to the central router to tear down the optical path.
- The central router sends release commands to both optical routers of source and destination node's optical routers to reset the MR's state to *"OFF"*.

4.3 Wavelength Allocation

For simplicity we use two static wavelength allocation mechanisms for OREX.

4.3.1 Single Data Stream Wavelength Allocation

In this case all available wavelengths within a waveguide are allocated to a single source-destination data stream(Figure 4.5-a), allowing a high data rate transfer when the optical path for communication is set for that pair. When network is congested however, a slow down in latency can occur due to multiple source-destination pairs requesting common paths for communication. To improve the OREX performance under heavy traffic loads we may consider of using multiple waveguides to provide more paths in the optical ring at the cost of increasing OREX with extra hardware cost (waveguides and the necessary MRs).

4.3.2 Multiple Data Stream Wavelength Allocation

Another possibility of improving the available path without increasing the waveguides is to use within the same waveguide a single or a group of wavelength as optical path. This increases the available path for OREX at the cost of lowering the bandwidth and increases the number of required MRs. Figure 4.5-b shows the case of two paths within the same waveguide using each 32-wavelength-channel.

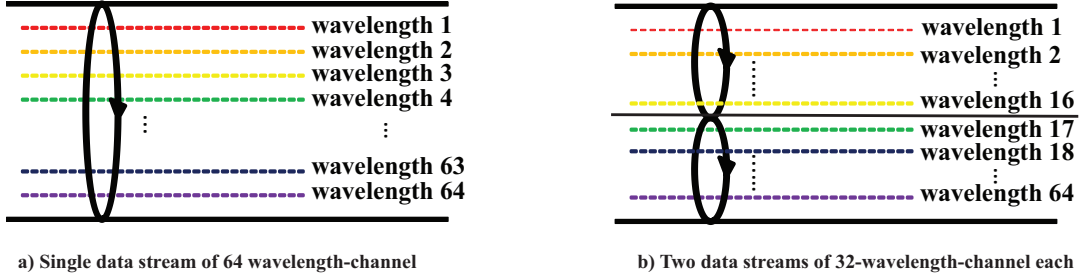


Figure 4.5: Single and two data streams waveguides

4.4 Cost Comparison

In this section we evaluate the hardware cost, power consumption cost, and achieved bandwidth for 64 nodes OREX and a hybrid photonic torus (HPTNoC) networks. The HPTNoC is also an hybrid NoC of electrical and optical networks. HPTNoC consists of 2 layers (optical and electrical) both using a planar torus topology to connect the network nodes.

4.4.1 Hardware cost

In term of hardware cost OREX presents a slightly better optical network compare to the HPTNoC.

The Table 4.4.1 summarizes the optical hardware count to build a 64 node OREX and HPTNoC. For the electrical component, OREX uses a single high radix crossbar of 64 input/output ports to connect the 64 nodes. The HPTNoC however uses 64 routers of 5 input/output ports to connect the 64 nodes using a torus topology. Both OREX and hybrid torus need 64 optical routers to connect the 64 nodes but the optical router of the OREX (2×2 input/outputs switch) requires only 4 MRs per node for the bidirectional ring instead of 12 for the HPTNoC's optical router (5×5

input/outputs switch).

64 Nodes Network	Hybrid Torus	OREX
Electrical Components		
Number of Port/Router	5	64
Number of Router	64	1
Total input/output ports	320	64
Optical Components		
Number of switches	64	64
Number of MR/Switch	12	4
Total MR	768	256

Table 4.1: Hardware cost comparison

4.4.2 Power Consumption

In this section we use the Phoenixsim [3] simulator to compare a 64 node OREX and a HPTNoC networks in terms of power consumption and bandwidth.

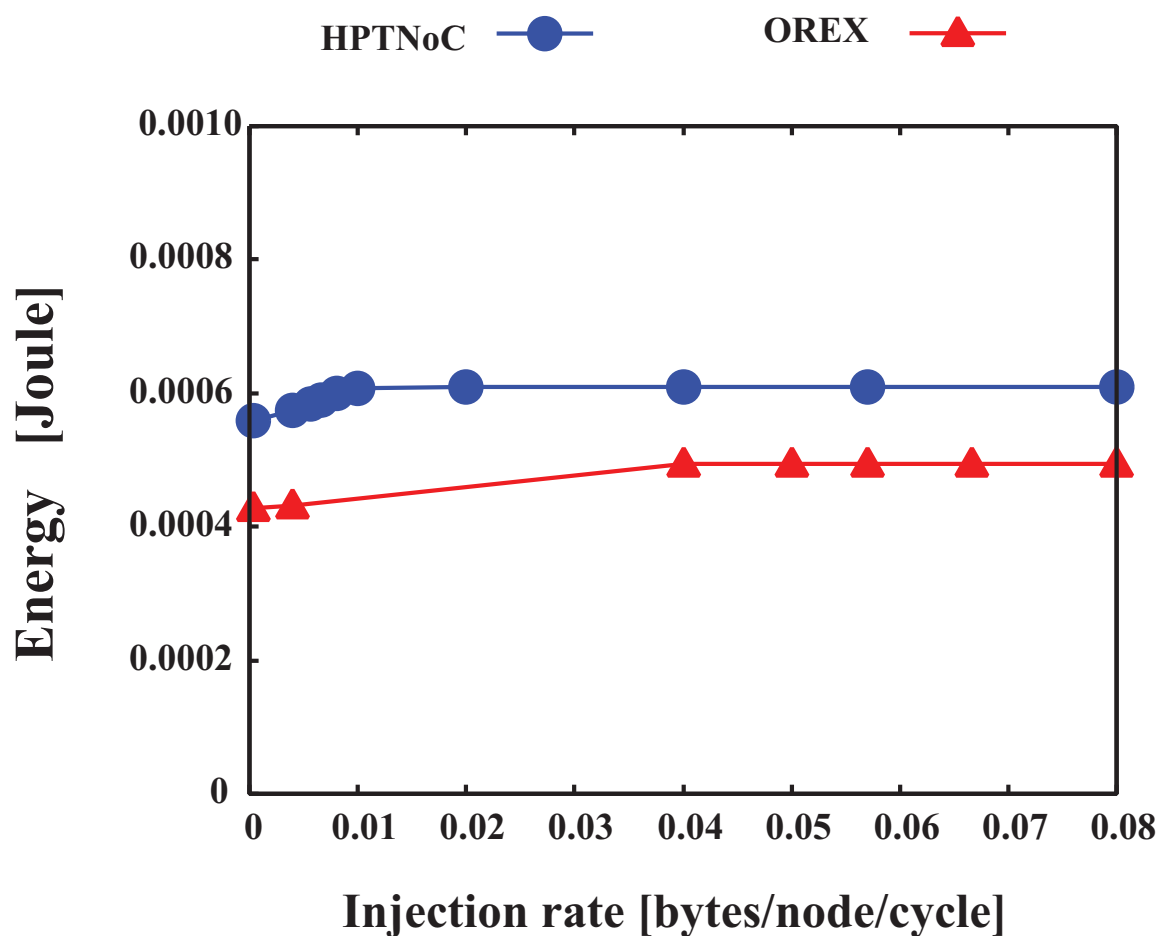


Figure 4.6: Power consumption comparison

Figure 4.6 shows that the OREX network consumes nearly 20 % less power than the hybrid torus due to less electrical and optical components.

4.4.3 Bandwidth

Figure 4.7 shows the achieved bandwidth for OREX and a hybrid torus networks. OREX achieved nearly 4 times the bandwidth of a hybrid torus network. By reduc-

ing the control network to a single central router, OREX manages to reduce the path setup time, hence achieves better bandwidth.

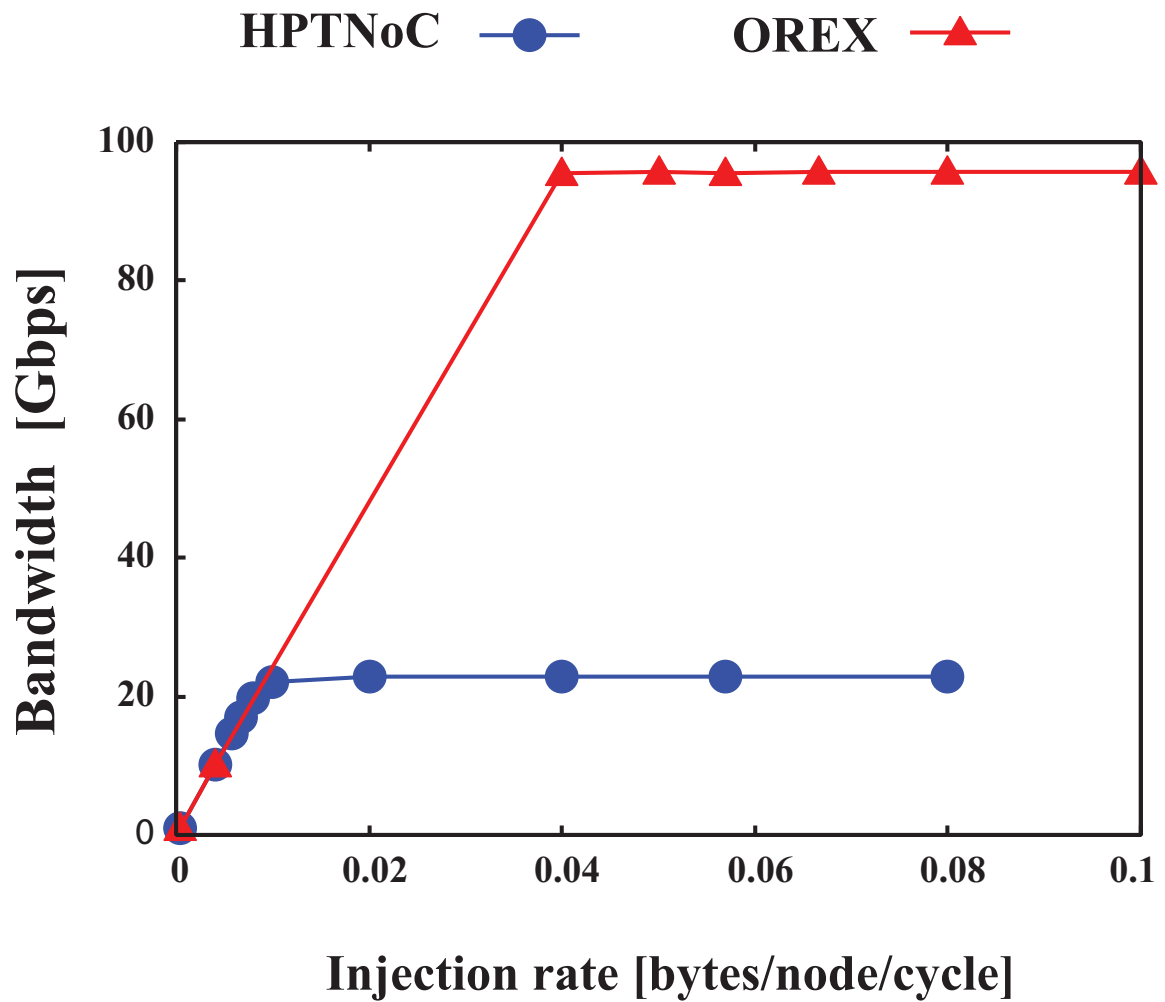


Figure 4.7: Bandwidth comparison

4.5 Performance Evaluation

Networks	Hybrid Torus, OREX
Size	64 Nodes
Router Frequency	5 GHz
Date rate /wavelength	12 GHz
# wavelength/waveguide	64
Total bandwidth/waveguide	800 Gbps

Table 4.2: Simulation parameters

We modified a network simulator *booksim* which was used in [6] to support OREX, and conducted experiments.

4.5.1 Simulation Conditions

We evaluated communication latency on OREX utilizing the network simulator. Our experimental simulation conditions are as follows;

[Network size]: 64 nodes are connected with the same number of optical routers.

[Sending overhead]: 5 cycles of the electrical central router are required for the setup processes 1 to 3 which are shown in figure 4.3 when there is no path conflict. As shown in the time diagram of the figure 4.8 setup request from source node need one cycle of link traversal (LT) to reach the central router. The request is decoded at the central router for routing computation (RC). The optical allocation stage (OA) in which request message compete for available optical path follows once the RC decide the path for the setup request. After Optical path allocation, the central router sends commands to turn on

the MR's of source-destination pairs which won the optical path competition setting the Optical switch (OS). In the same time acknowledgment packet of path establishment performs virtual channel and switch allocation (VSA) then traverses the link between the central router and the node (LT).

[Optical path allocation] : Allocate a router cycle ² necessary for optical data transfer. In another word, size of optical data transfer is $signal_rate \times cycle_time \times \#wavelengths$.

[Release process] : Same as setup request, the release request also requires 5 cycles to tear down an optical path. The time diagram is shown in Figure 4.8.

[Wavelength per waveguide] : We use 64 wavelengths per waveguide.

[Wavelength assignment] : As we described in section 4.3, we tested two cases: single, and multiple data stream wavelength allocation mechanisms. The path multiplicity is implemented using single or multiple waveguides.

- single waveguide: in this case a single waveguide is used for optical data transfer. Path multiplicity is allowed by dividing the 64 available wavelengths within the waveguide into group of data streams: one(64 wavelengths/path), two (32 wavelengths/path), four (16 wavelengths/path), and eight (8 wavelengths/ path) ring cases are evaluated.
- multiple waveguide: in this case we use multiple waveguides for path multiplicity. single, double, and quadruple waveguides in which one, two and four waveguides are used. The table 4.5.1 details the different simulated OREX network configurations.

[Traffic pattern] : Uniform-random and neighbor-to-neighbor traffic

²time necessary for achieving a single pipeline stage (e.g. RC or LT).

Network	Description
1-ring, 1-waveguide (single optical path, an optical path in this case consist of 64 wavelengths).	OREX that uses a single waveguide, all wavelengths are allocated to a single data stream.
2-rings (2 optical paths, an optical path in this case consist of 32 wavelengths)	OREX that uses a single waveguide, wavelengths within the waveguide are divided in 2 groups of 32 wavelengths each for data stream.
4-rings (4 optical paths, an optical path in this case consist of 16 wavelengths)	OREX that uses a single waveguide, wavelengths within the waveguide are divided in 4 groups of 16 wavelengths each for data stream.
8-rings (8 optical paths, an optical path in this case consist of 8 wavelengths)	OREX that uses a single waveguide, wavelengths within the waveguide are divided in 8 groups of 8 wavelengths each for data stream.
2-waveguides (2 optical paths, an optical path in this case consist of 64 wavelengths)	OREX with 2 waveguides, each waveguide is allocated for a single data stream.
4-waveguides (4 optical paths, an optical path in this case consist of 64 wavelengths)	OREX with 4 waveguides, each waveguide is allocated for a single data stream.

Table 4.3: OREX simulations configurations

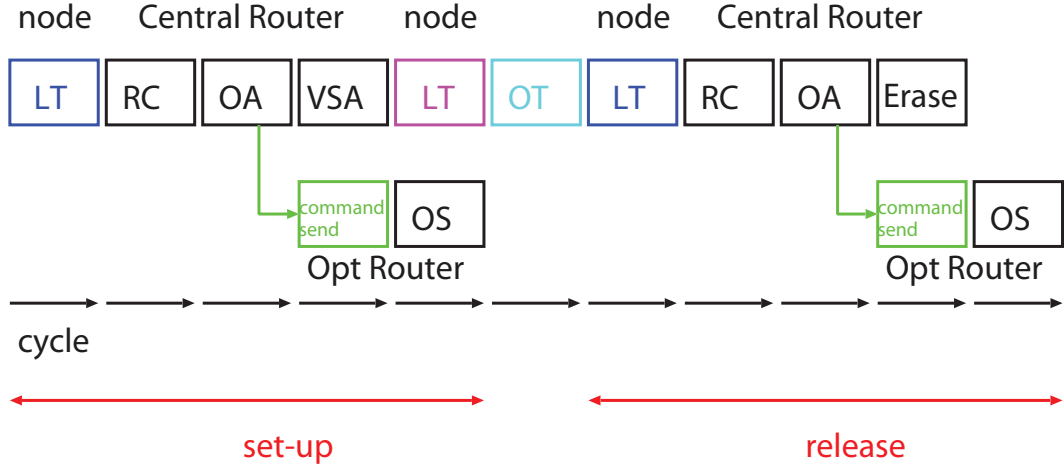


Figure 4.8: A time diagram of communication on OREX.

4.5.2 Experimental Results

4.5.2.1 Zero-load latency

Zero-load latency gives a lower bound on the average latency of a packet through the network under the assumption that a packet never contends for network resources with other packets. The Figure 4.9 shows the zero-load latency of OREX for one, two, four and eight rings when using a single waveguide. The results confirm that under low injection rate allocating all the available wavelengths to a single message stream present a better solution compared to dividing them among multiple data streams. The division of the available wavelengths among several data stream reduces the available bandwidth for the optical data transfer. Hence one-ring outperforms two, four and eight rings.

4.5.2.2 Latency Evaluation under uniform and neighbor traffic patterns

The Figure 4.10 shows that using multiple paths improves the performance of OREX for uniform and neighbor traffic patterns.

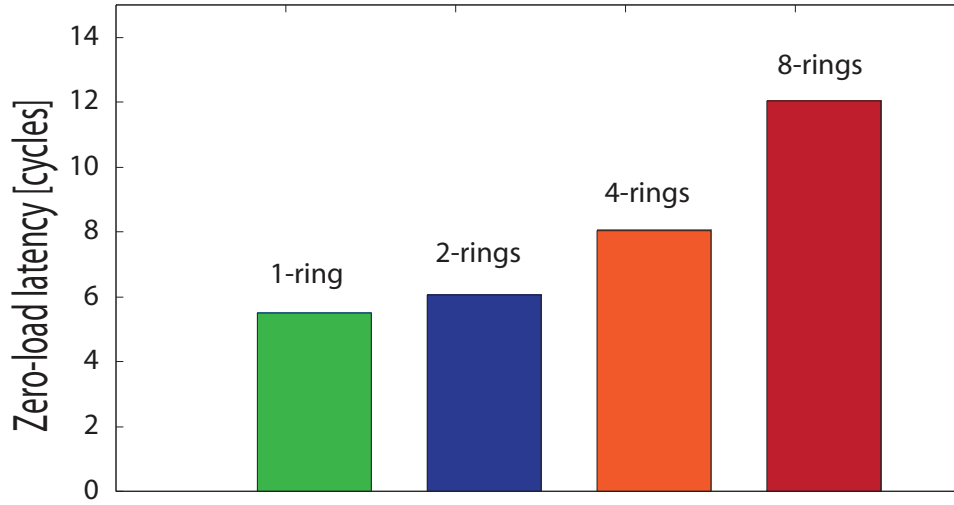


Figure 4.9: Zero load latency.

For uniform traffic pattern, when we divide the available wavelengths within a single waveguide among multiple message streams as we see in Figure 4.10a, more group of data stream there are higher is the saturation load of the OREX. 1-ring OREX is saturated at the lowest traffic load. Although the OREX communication bandwidth is reduced when using multiple data streams, at higher traffic load the contention is reduced with the availability of multiple optical paths(2, 4, and 8 rings). The multiple-rings OREX hence can load higher traffic before saturation. To maintain the same bandwidth when using multiple paths, multiple waveguide are used as show in Figure 4.10c. Multiple waveguides further improves the performance of OREX.

In neighbor traffic patterns, nodes communicate with their neighboring nodes, a slow down in performance can occurs when using path multiplicity that only reduce the bandwidth for data stream. As show in Figure 4.10c, using 8-rings data streams, the performance of the network is considerably reduced by the decrease in bandwidth. Thus a path multiplicity of 8 is not required under this traffic load for a 64 nodes OREX.

Chapter4. OREX: Hybrid Optical Ring Electrical Crossbar Network-on-Chip

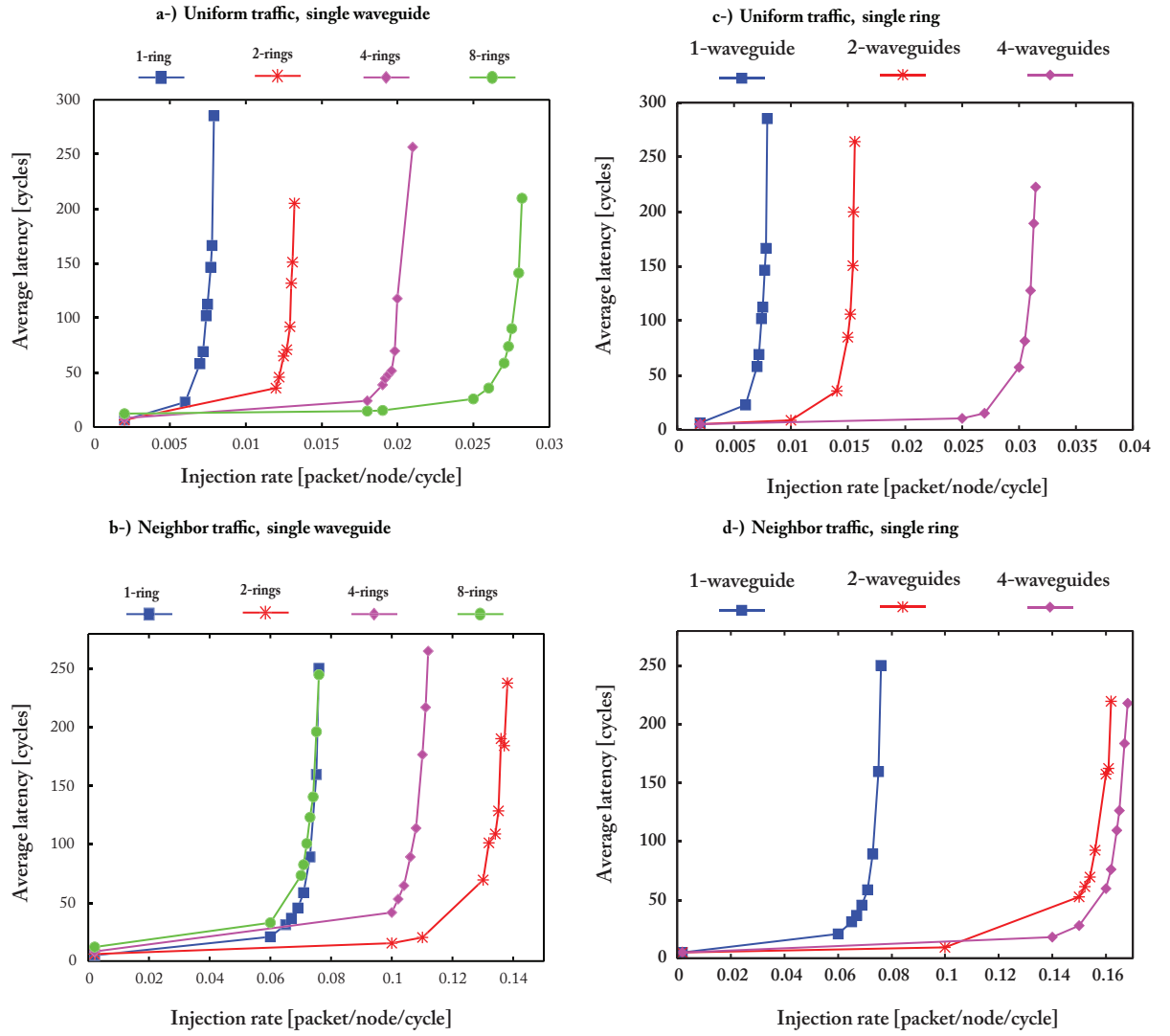


Figure 4.10: Average latency of OREX under uniform (a , c), and neighbor(b, d) traffic patterns.

Chapter4. OREX: Hybrid Optical Ring Electrical Crossbar Network-on-Chip

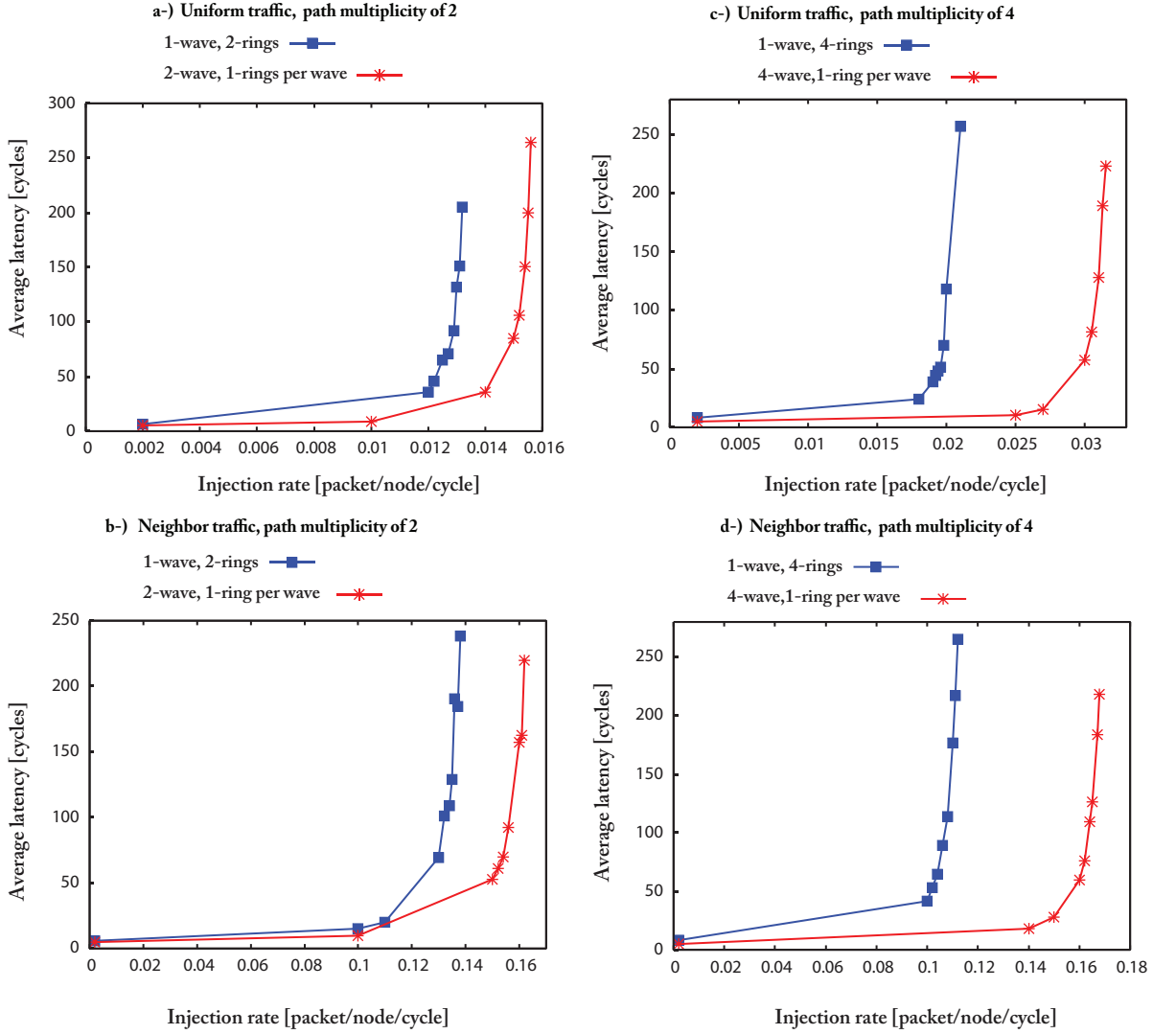


Figure 4.11: Effect of path multiplicity using single and multiple waveguides under uniform (a, c) and neighbor (b, d) traffic patterns.

The Figure 4.11 shows the effect of path multiplicity using multiple waveguides and a single waveguide with division of the wavelengths among multiple data streams for uniform and neighbor traffic patterns. Results show that, the path multiplicity using multiple waveguides achieves better throughput compare to the single waveguide with multiple ring. The improvement in performance is however

at the cost higher hardware cost (more waveguides).

4.6 Conclusion

In this Chapter, we have proposed OREX, which is a hybrid NoC consisting of an optical ring and an electrical central router. OREX is designed using high-performance and low-power NoC by integrating nanophotonic technology as well as a traditional electrical indirect network. We evaluated the OREX performance using static wavelength allocation under probabilistic traffic patterns and show that OREX present better performance and power consumption compare to a hybrid Torus network. In the next chapter a fully optical NoC is proposed which will take full advantage of optical interconnect without the need of electrical control network to further reduce power consumption cost. The proposed fully optical NoC further improves performance by integrating static and dynamic wavelength allocations with selection mechanisms.

CHAPTER 5

FULLY OPTICAL RING NETWORK-ON-CHIP

In this Chapter a new optical NoC that addresses the following issues of previous works is presented:

- Hybrid NoCs suffer from high power consumption of either electronic path setup networks [12, 25, 4] or local communication using electrical interconnect [13, 18, 20].
- Previously proposed photonic NoCs in one hand use only low bandwidth static wavelength allocation [18, 2, 14, 27], with short or without arbitration overhead. On the other hand, photonic NoCs with only high bandwidth dynamic wavelength allocation suffer of higher arbitration overhead [25].

The architecture has the advantage of being a fully optical hence low power NoC, that can employ static and dynamic wavelength allocation techniques in the same network. It consists of optical switches connected using three waveguides in a

multi-ring topology. These rings of waveguides are used for static and dynamic wavelength allocation communications, and arbitration respectively. The architecture takes advantage of both wavelength allocation mechanisms by selecting the adequate one depending on communication message sizes (baseline selection) or congestion informations (contention based and smart selection).

5.1 A Fully Optical Ring NoC's Architecture

Figure 5.1 illustrates the general overview of our proposed fully optical ring NoC (FORNoC) for a network of 8 nodes. It consists of a Laser source, and three waveguides that connect the nodes using a ring topology. The first waveguide is used for static communication, the second one dynamic communication, and the third as arbitration waveguide. The arbitration waveguide consists of the same number of wavelength-channels with the number of nodes. A token is assigned to every wavelength-channel, each representing the right to modulate optical data intended for a particular node.

In the static communication waveguide, a single wavelength-channel is statically allocated for each destination node as receiving channel. The destination node receives optical data from a sender node by switching “ON” the detector of the wavelength-channel uniquely assigned for that particular node.

The dynamic communication waveguide consists of multiple wavelengths which are shared by all nodes. Unlike the static waveguide, wavelengths are dynamically allocated by a manager node to source-destination communication pairs. The manager node is a special node, denoted N0 in Figure 5.1. It performs dynamic wavelength allocation based on requests in execution time.

Figures 5.2(a) and (b) show microarchitecture of the normal and manager nodes, respectively. The normal node consists of electronic input and output buffers, arrays of modulators / detectors (silicon photonic devices), and a controller. The controller

is used for switching state of the modulators and detectors to modulate / detect optical data stream into / from a waveguide. In addition, the manager node contains a wavelength allocator.

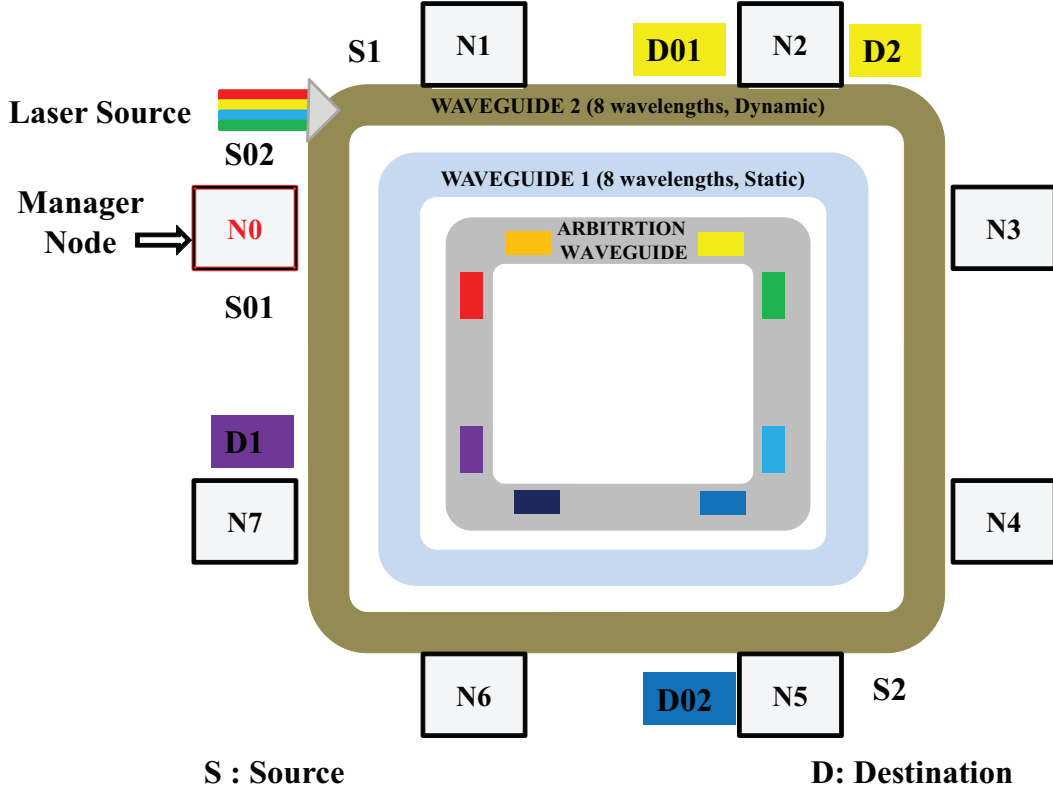


Figure 5.1: FORNoC architecture

5.2 Communication Mechanisms

Our proposed architecture offers two types of communications: static and dynamic. The static communication is based on a token-based arbitration. The dynamic communication uses a manager node to allocate wavelengths to source-destination communication pairs. While static communication requires a low communication overhead, it offers only a single wavelength-channel bandwidth for data transfer. The

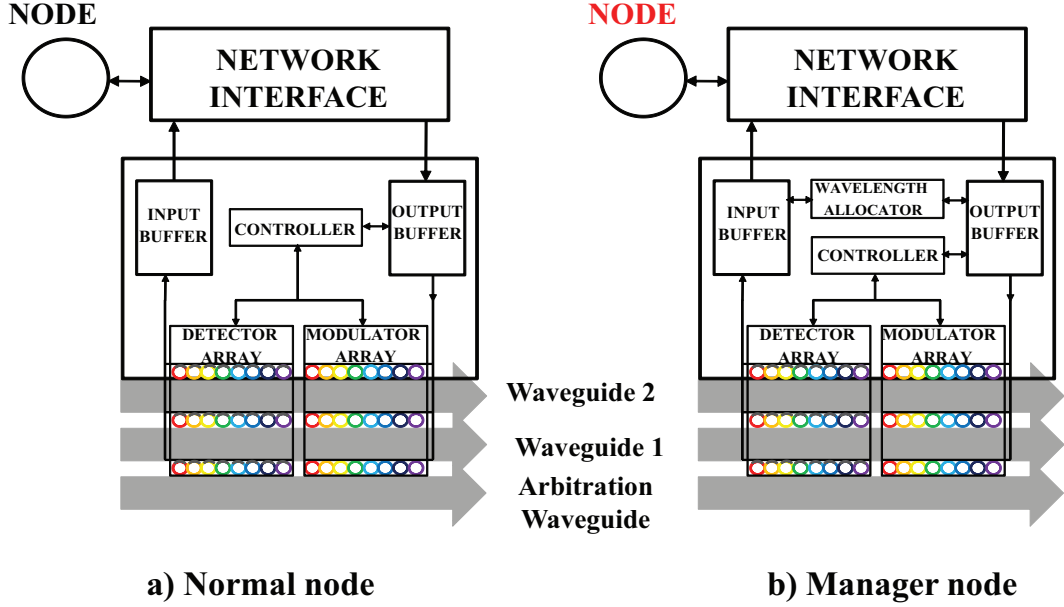


Figure 5.2: Nodes microarchitecture.

dynamic communication on the other hand offers higher bandwidth at the cost of a higher arbitration overhead of requesting wavelength allocation to the manager node.

5.2.1 Static Communication

Figure 5.3 shows the nodes connection in the static communication waveguide. Each network node can read from only its dedicated receiving wavelength-channel and can write to any over node's receiving wavelength-channel. Contention of multiple source nodes to the same destination node is resolved using token ring arbitration. Static communication has the advantage of low communication establishment overhead, however its bandwidth is limited to a single wavelength-channel.

Let's consider a static communication between node N1 (as source S1) and node N7 (as destination D1) shown in Figure 5.1. By following the communication steps of Figure 5.4 which shows the pipeline stages of a static communication, node N1 injects an electronic message data, to the network interface, which is saved in the

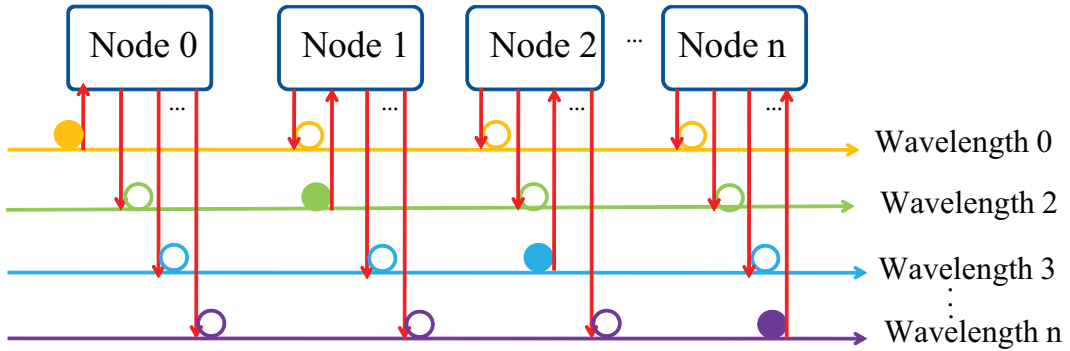
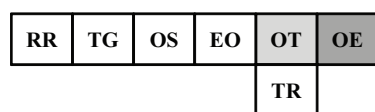


Figure 5.3: Nodes connection in the static communication waveguide.

node's output buffer. The controller reads its destination address (node N7) from the message header (RR). Next, a detector, associated to the wavelength for the destination node N7, is switched "ON" to grab the token for sending data on node N7's specified receiving wavelength-channel (TG). When source node N1 grabs the token, it sets up related modulator (OS) to prepare the optical data modulation. Electrical message data are modulated into optical data (EO) by node N1 and injected onto the static waveguide (node N7's receiving wavelength-channel). Then, modulated optical data are transferred on the statically assigned destination node N7's receiving wavelength-channel (OT), and finally the grabbed token is released by the source node N1 (TR) when data modulation is completed. Destination node N7 detects the optical data transferred on the static waveguide and converts them into electronic data (OE). Note that each pipeline stage of the Figure 5.4 may take multiple cycles depending on the message size and the token availability (congestion).

5.2.2 Dynamic Communication

Let's consider a dynamic communication between node N5 (as source S2) and node N7 (as destination D2) shown in Figure 5.1. A dynamic communication is a com-



RR: Read Request

TG: Token Grant

OS: Optical Switching Setting

EO: Electrical to Optical Conversion

OE: Optical to Electrical Conversion

TR: Token Release

OT: Optical Traversal

Figure 5.4: Time diagram of a static communication

combination of static communications (steps 1 and 3), a wavelength allocation (step 2), and data transfer (steps 4 and 5) shown in the time diagram of Figure 5.5. The dynamic communication can be divided into two phases as in a circuit switching communication: path setup (steps 1 to 3) and data transfer (step 4 and 5). First the source node N5 sends a request to the manager node N0 (Step 1, static communication in which N5 and N0 are the source and destination, respectively). When the manager node N0 receives node N5's request and there is a free path in the dynamic waveguide between source node N5 and destination node N7, the manager node allocates the path for the pair (step 2) and sends grant messages using static communication to both N5 and N7, source and destination nodes, respectively (step 3). It's important to notice that in this step, node N7 and node N5's tokens for static communication may not be available at the same time, however the grants are sent only when both tokens are grabbed by N0. After N5 and N7 nodes receive the path grant messages sent by manager node N0 (step 4), the source node N5 modulates the data to the dynamic waveguide for data transfer. Destination node N7 detects the data on the dynamic waveguide and the communication ends with a tear down message (step 5).

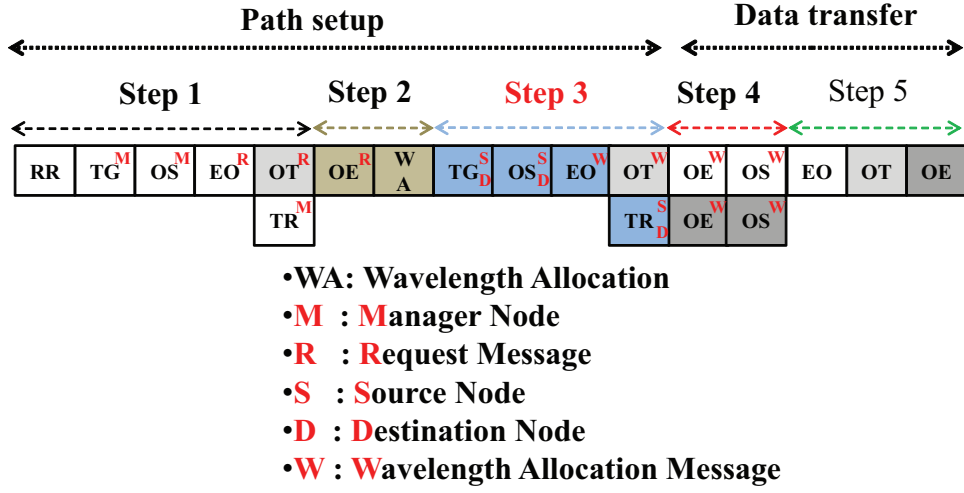


Figure 5.5: Time diagram of a dynamic communication

5.2.3 Bended Static and Dynamic Communications

Both static and dynamic communications may occur at the same time, in a bended way. Let's assume that, the previous communication examples of Sections 5.2.1, and 5.2.2 happen at the same time. In this case, both static and dynamic communications have the same destination node. The static communication between source node N1 and destination node N7 uses node N7's receiving wavelength-channel of the static communication waveguide. In the step 3 of the dynamic communication between source node N5 and destination node N7, the manager node N0 has to send the path grant message to destination node N7. Hence, it also need to use the same node N7's receiving wavelength-channel. If the static communication between N1 and N7 is still not completed when the dynamic communication between node N5 and N7 reaches the step 3; as the token for N7's receiving wavelength channel is not available, the manager node N0 will delay the following step of the dynamic communication until the static communication between source N1 and destination N7 finishes. As in this example, the token based arbitration of static communication,

and the manager node in dynamic communication help to solve any contention that may take place.

5.3 Wavelength Allocation Selection Mechanisms

A key point of our architecture is the possibility to choose between two wavelength allocation mechanisms. On one hand static allocation offers a quick establishment of communication between nodes with low data transfer bandwidth. On the other hand, dynamic allocation with high bandwidth, suffers from higher overhead of communication establishment. In this Section, we describe how we take advantage of both communication mechanisms to achieve good performance.

5.3.1 Baseline Selection Mechanism

Let's assume Lat_{static} , and $Lat_{dynamic}$, the zero-load latencies for sending a message using the static and the dynamic allocation mechanisms, respectively.

Lat_{static} can be defined by Equation (5.1) as:

$$Lat_{static} = Lat_{setup static} + \frac{message_{size}}{BW_{static}} \quad (5.1)$$

where $Lat_{setup static}$ is the latency for path setup, and BW_{static} is the bandwidth for the static allocation mechanism.

$Lat_{dynamic}$ can be defined by Equation (5.2) as:

$$Lat_{dynamic} = Lat_{setup dynamic} + \frac{message_{size}}{BW_{dynamic}} \quad (5.2)$$

where $Lat_{setup dynamic}$, and $BW_{dynamic}$ are the path setup latency, and the bandwidth for the dynamic allocation mechanism, respectively. Although latency of static communication seems larger than the latency of dynamic communication this situation may change for certain message sizes. Because of the low bandwidth of data transfer

in static communication, the data transfer time can be considerably high for large message sizes. The higher overhead of path setup in dynamic communication will no longer be a disadvantage for such cases, because of its higher data transfer bandwidth. Let's assume a communication case in which: i) both latency of static and dynamic communications are equal for a given message size; ii) a single wavelength-channel bandwidth is used for static communication while n ($n > 1$) wavelengths are used for dynamic communication, we can derive Equation (5.3):

$$message_{size} = \frac{Setup_{diff} \times n \times BW_{static}}{(n - 1)} = threshold \quad (5.3)$$

as $BW_{dynamic} = n \times BW_{static}$ and $Setup_{diff}$ is the setup time difference between static and dynamic communications. Equation (5.3) defines the threshold message size for which static communication outperforms the dynamic communication. For any value of message size higher than *threshold*, the dynamic communication outperforms the static communication. Using this threshold, we can classify messages as small or large. The normal selection mechanism selects between the two communication modes using the message size. While static allocation mechanism is selected for message sizes smaller than the threshold, the dynamic allocation mechanism is selected for higher message sizes.

5.3.2 Contention Based Selection Mechanism

When the network is highly loaded, the latency for dynamic communication quickly increases and many dynamic communication requests have to wait for resource allocation. Under such situation, there is a trade-off between waiting for high bandwidth dynamic communication resource to be freed, and a quick establishment of low-bandwidth static communication. In order to optimize the utilization of both static and dynamic communications, we introduce a smart selection mechanism that helps to choose static or dynamic communication under the congested situations.

Manager node checks the number of waiting request messages for dynamic communication to confirm congestion. The congestion status is defined based on a threshold number of waiting request messages in the manager node. The smart selection mechanism refuses further dynamic communication requests when this threshold is reached and notifies the requester source nodes to select static communication rather than waiting a long time for the dynamic resource. We can expect that, this mechanism alleviates congestion in the dynamic communication and improves performance. Experimental results are shown in the next Section.

5.3.3 Smart Selection Mechanism

Figure 5.6 describes the smart selection mechanism. It uses network information to adaptively allocate communication bandwidth to requester source nodes. Depending on message size a source node will request $n\text{-wavelength} - \text{channel}$ for dynamic communication to the manager node. The manager node in return will allocate the requesting number of wavelength channel if available. If the requested number of wavelength are not available, the manager node will lookup for, half or quarter or one-eighth the number of desired wavelength depending on their availability, respectively. By dynamically allocating a bandwidth on communication based, the smart allocation fully take advantage of the dynamic communication's waveguide.

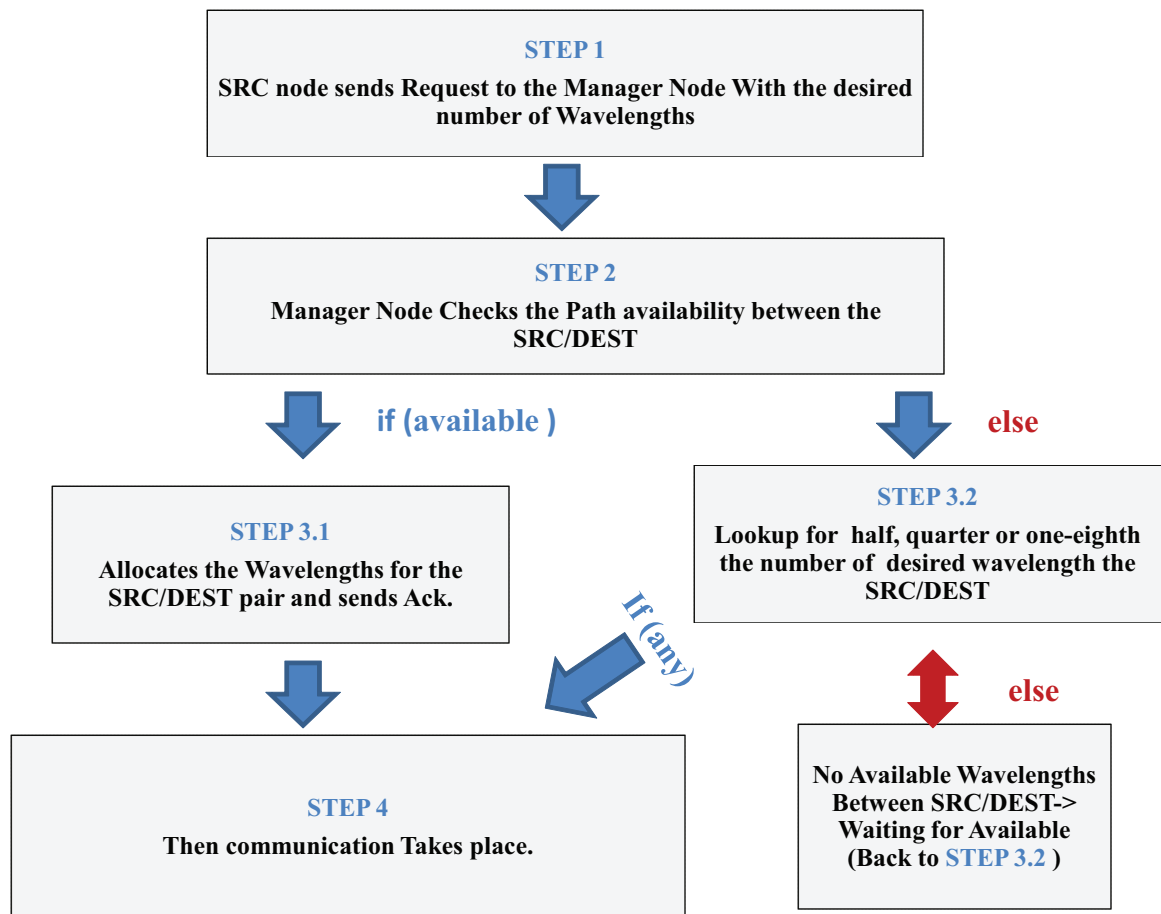


Figure 5.6: Smart selection mechanism.

CHAPTER 6

SIMULATION RESULTS AND ANALYSIS

This chapter discusses the performance of the fully optical ring NoC (FORNoC). After a brief comparison with several optical NoCs in terms of hardware cost, we use a modified version of PhoenixSim [3] photonic NoC simulator to evaluate performance and energy consumption of FORNoC for different probabilistic traffic patterns.

6.1 Hardware Cost Comparison

To build a 64-node network, FORNoC uses a total of 16 256 ring resonators (254 per node, 126 for static waveguide, 126 for arbitration waveguide, and 2 for the dynamic waveguide), 3 waveguides (static, dynamic, and arbitration), 8256 photodetectors (127 per node, 63 for static waveguide, 63 for arbitration waveguide, and 1 for dynamic waveguide). Table 6.1 summarizes the hardware cost requirements

of a 64-node for a hybrid 2-D planar (mesh, torus) networks [25], PROPEL [18], OREX [4], CORONA [28], and FORNoC. Compared to hybrid architectures, fully optical FORNoC and CORONA networks do not require electronic switches. The fully optical networks however use more optical components necessary for arbitration. FORNoC uses fewer optical waveguides and ring resonators compared to CORONA by providing both static and dynamic wavelength allocation techniques. CORONA however uses fewer photo detectors (less dedicated paths).

6.2 Simulation Setup

Following are some simulation setups we use to evaluate the performance of our architecture:

[Network size and wavelengths]: As FORNoC uses a single wavelength-channel per node for the static communication, the number of required wavelengths for static communication is proportional to the number of nodes in the network. The same number of wavelength is also required for the arbitration. Thus to implement 32-node network, we use two waveguides with each of them using only 32 wavelengths per waveguide for static, and arbitration waveguides, respectively and a third waveguide with 64 wavelengths for dynamic communication. As most of previous works suggest using a maximum of 64 wavelengths per waveguide, for the case of 128 nodes, we used two waveguides of 64 wavelengths each to connect the node statically as well as to perform arbitration; meanwhile a single waveguide of 64 wavelengths is used for the dynamic communication. Hence our architecture requires a total of 5 waveguides to implement 128 nodes.

[OREX]: For fair comparison with FORNoC, we use the same number of communication waveguides for both architectures. While FORNoC uses static and

Table 6.1: Architecture hardware cost comparison for 64-node networks.

	2-D Hybrid (Mesh, Torus) [25]	PROPEL [18]	OREX	CORONA [28]	FORNoC
Wavelengths	64	64	64	64	64
Waveguides	64	64	2	64	3
Ring Resonators	1024	3072	256	72192	16256
Photodetectors	4096	1536	8192	7424	8128
Electrical Switches	5 × 5 (64)	5 × 5 (16)	64 × 64 (1)	-	-

Chapter6. Simulation Results and Analysis

dynamic communication waveguides, OREX uses both waveguides (clockwise and counter-clockwise) for dynamic communication.

[PMNoC]: The PMNoC is a hybrid NoC proposed in [25] with mesh topology. The network is formed by a mesh optical NoC overlaid by a similar mesh electronic path setup network.

[Measurement]: The communication latency is measured as the time to transfer the whole message, from when it is created to when the message reaches its destination. We evaluate the average latency and average bandwidth of the networks as a function of the message injection rate during a simulation time. The average network latency/bandwidth, for an injection rate is depicted as the average latency/bandwidth of all messages that reach their destinations during the simulation time.

[Clock Frequency and Speed of Modulation]: For the clock frequency, we use 5GHz as used in [25, 20, 28]. Although 12.5Gbps [20], 40Gbps [25] modulation speeds have been suggested, we use 10Gbps as in [27, 18] for our simulations.

[Congestion based selection threshold]: The smart selection threshold depends on several experimental parameters. In the experimental conditions of this work, based on simulations, we use 35 waiting dynamic communication requests for the congestion based selection threshold. Further requests for dynamic communication are directed to use the static communication.

[Message Size]: We use five different message sizes. Based on consideration in Section 5.3.1, 12 and 20 bytes for small size messages; and 256, 400, 516, and 1024 Bytes for large size messages. Messages of different sizes are randomly generated with either the same probability (SP) or different probability (DP). Depending on selection mechanisms, messages will be allocated static or dynamic communication.

Chapter6. Simulation Results and Analysis

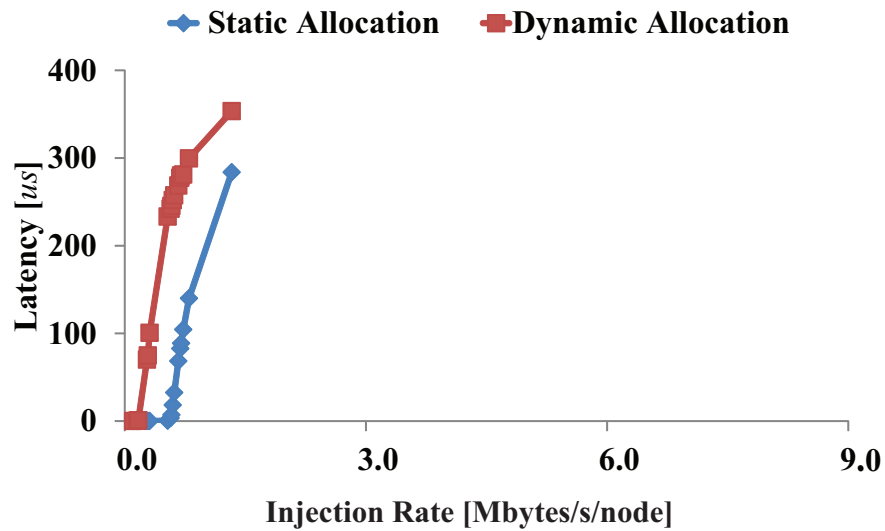
[Smart selection]: In the case of smart selection, 32-*wavelength-channel* bandwidth is set as desired bandwidth for 1024 Bytes, 16-*wavelength-channel* for 512 Bytes, and 8-*wavelength-channel* for 256 Bytes. The manager node will allocate the desired bandwidth when available, and adapts the communication bandwidth to half, quarter or one-eighth of the desired bandwidth in case of congestion.

Table 6.2 summarizes our simulation parameters.

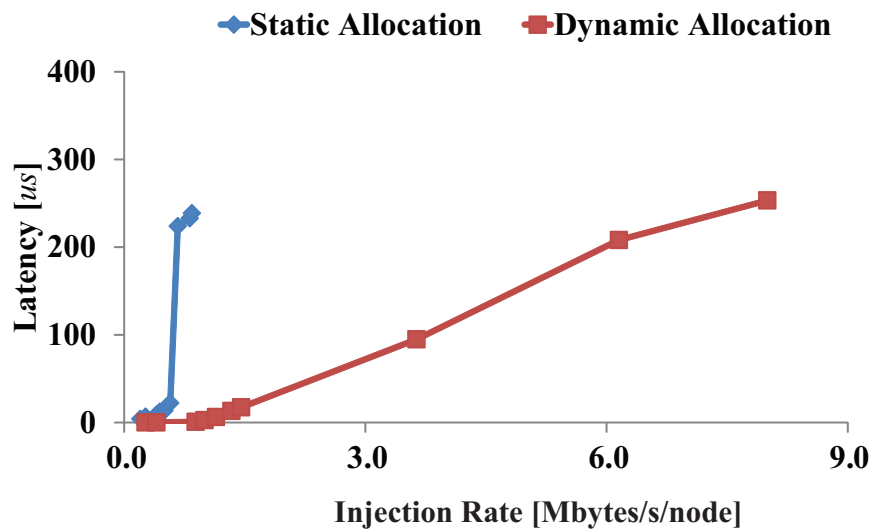
Table 6.2: Simulation parameters

Parameter	Setting
NoC Architecture	FORNoC, OREX, PMNoC
Number of Nodes	32, 64, 128
Traffic Patterns	Uniform, neighbor, hotspot
Message Sizes	12, 20, 256, 400, 516 and 1024 Bytes
Communication Channel # wavelength ×#waveguide	32×2, 64×2, 64×3
Clock Frequency	5GHz [25]
Speed of Modulation	10Gbps [27, 18]
Communication Types for FORNoC	Static only, dynamic only, and combination

6.3 Static and Dynamic Communications Comparison



a) 20 Bytes message size.



b) 400 Bytes message size.

Figure 6.1: Performance comparison of static and dynamic communications

In this Section, we compare the performance of the static and dynamic communications standing alone. Figure 6.1 shows the simulation results for static and dynamic communications under uniform traffic pattern. Static and dynamic communications are simulated for 20, and 400 Bytes message sizes, respectively. These results confirm our assumption in Section 5.3 that, for small message size (20 Bytes), the static communication outperforms the dynamic one. Because of the small message size, the fast path latency of the static communication is preferable in this case (Figure 6.1 (a)). On the other hand, for larger message size, higher bandwidth is more efficient than slower path setup. As shown in Figure 6.1 (b), dynamic communication performs better in such case.

6.4 Performance and Energy Consumption Comparison

In this Section we compare the FORNoC with baseline selection and smart selections, PMNoC, and OREX NoCs of 64 nodes under uniform traffic pattern. 12, 256, 512, and 1024 Bytes message sizes are randomly generated with the same probability (SP).

(1) Energy consumption comparison

In our simulations, design parameters such as static and dynamic energy of every component are integrated. Energy consumed for injection, ejection arbitration, buffering opto-electrical conversions, data transfer are calculated during simulation execution time. The main energy consumption difference between the hybrid NoCs (PMNoC, OREX) and FORNoC is the arbitration energy as the hybrid NoCs perform the arbitration electronically. For a given injection rate we depicted the average energy consumed in the network for PMNoC, OREX and FORNoCs. Figure 6.2 shows

Chapter 6. Simulation Results and Analysis

the average energy consumed versus the injection rate under a uniform random traffic pattern. The results show that PMNoC and OREX consume higher energy as compared to FORNoC networks. Both PMNoC and OREX consume higher energy by exchanging control messages between source and destination via electronic path setup networks whereas FORNoC performs those tasks optically. FORNoC with smart selection consumes nearly similar amount of energy as the FORNoC with baseline selection because only fewer overheads are added to the arbitration which in turn reduces the energy consumed with an improvement in path allocation. Because power constraint is so severe in future NoCs, FORNoC can be an alternative low power solution to the hybrid NoCs.

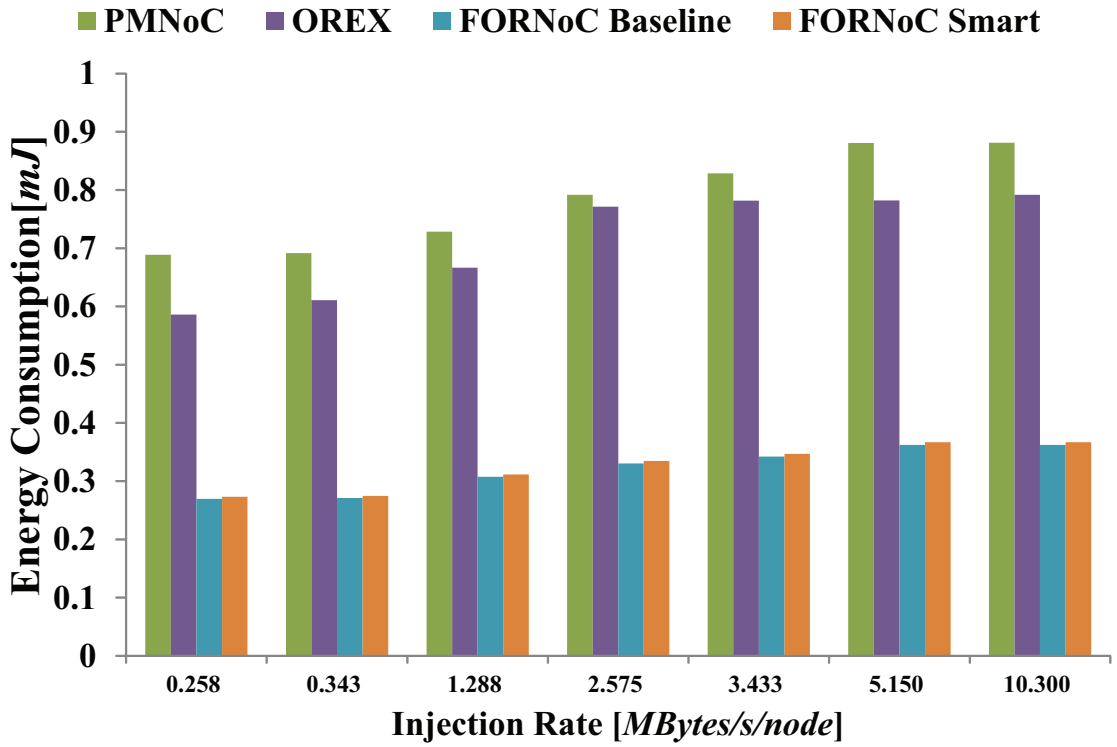


Figure 6.2: Energy consumption comparison.

(2) Latency and bandwidth comparison

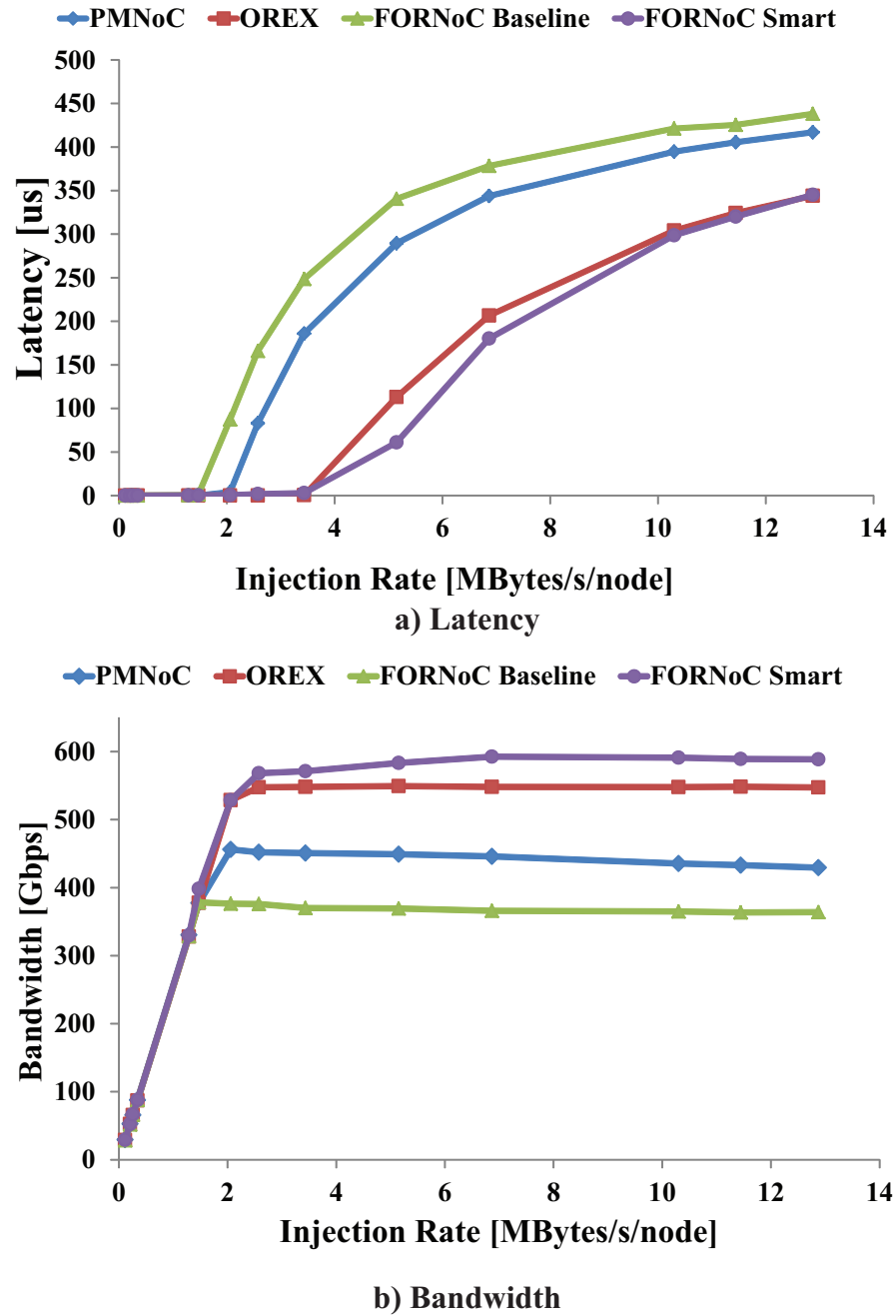


Figure 6.3: Latency and bandwidth performance.

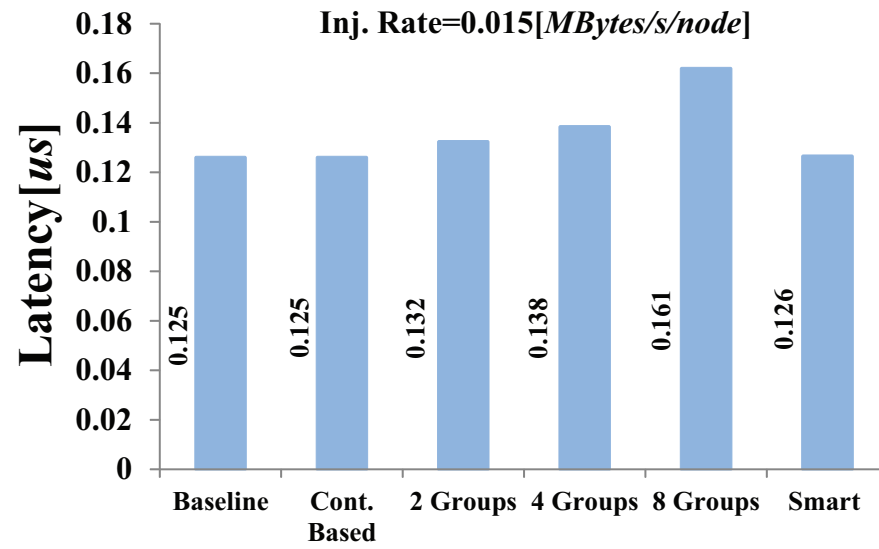
Figure 6.3 shows the performance of PMNoC, OREX and FORNoCs in terms of latency (a), and bandwidth (b) under uniform random traffic, respectively. OREX outperforms PMNoC and FORNoC with baseline selection mechanism in average latency and bandwidth. The OREX has a low latency path setup network (electronic crossbar) which explains its latency and bandwidth performances. By adapting the dynamic communication bandwidth depending on network utilization, the FORNoC with smart selection outperforms the other NoCs in term of latency and bandwidth.

6.5 FORNoC with different selection techniques

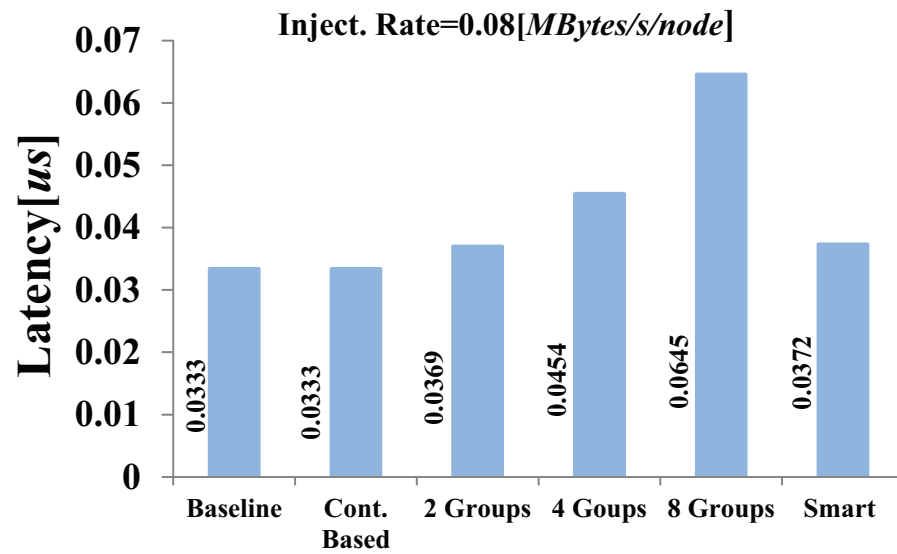
In this Section, we compare FORNoC performance using different wavelength allocation and selection techniques. Under a uniform traffic pattern, the baseline, contention based (Cont. Based), grouping and smart selection are compared. Different message sizes (12, 256, 512, and 1024 Bytes) are randomly generated with the same probability (SP) and different probability (DP). For DP, the 12 Bytes messages are generated with a probability of 5%, 256 Bytes with 15%; 512 Bytes with 30% and 1024 Bytes with a probability of 50%.

(1) Low load latency comparison

Figure 6.4 shows the performance of FORNoC configurations at very low load traffic (when almost no congestion occurs). For both SP and DP, the Baseline FORNoC outperforms all other configurations. Because no congestion occurs in the network, the baseline selection technique which select static communication for small message size and dynamic communication for large message sizes, provides the highest dynamic communication bandwidth (*64-wavelength-channel*). thus outperforming the other configurations.



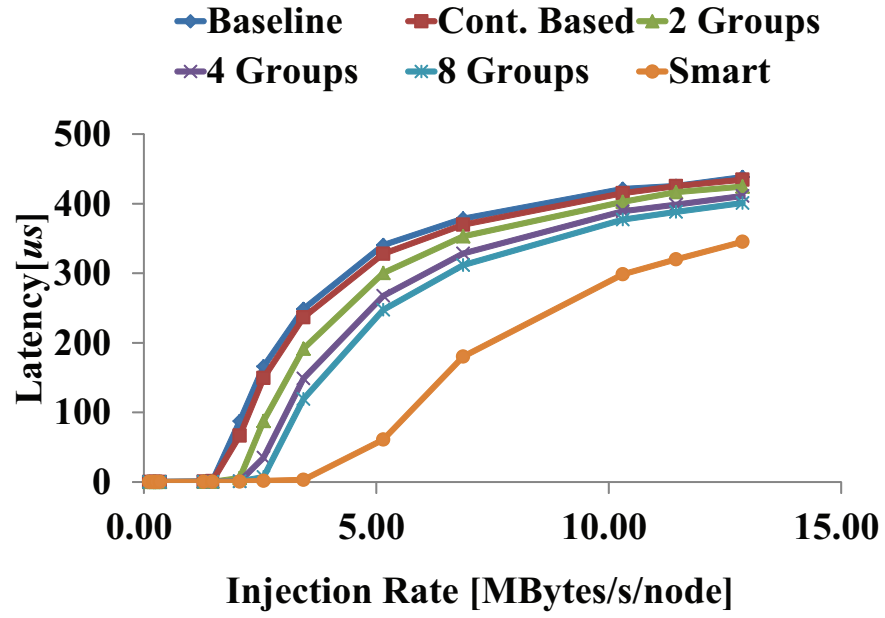
a) SP injection



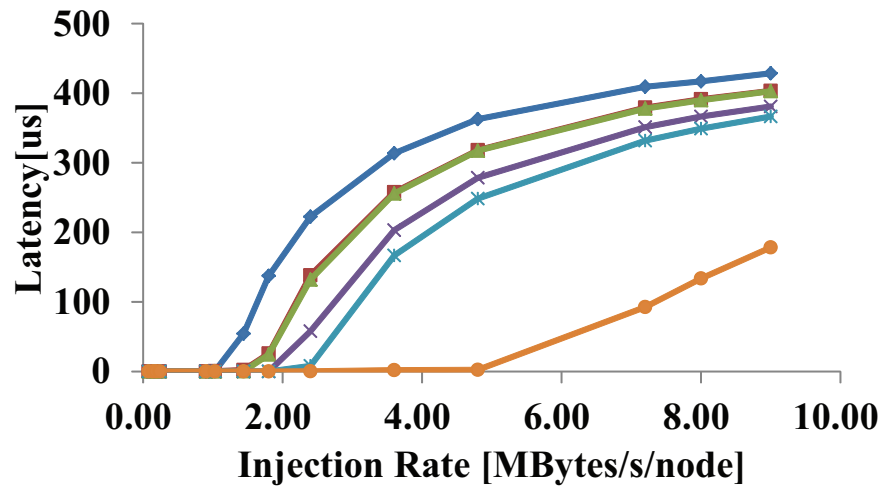
b) DP injection

Figure 6.4: Low load latency under uniform traffic (SP)

(2) latency performance comparison



a) SP injection



b) DP injection

Figure 6.5: Latency performance comparison of FORNoCs.

Figure 6.5 (a) and (b) show the result for same (SP) and different (DP) probability message injections, respectively. Results show that the FORNoC with smart selection outperforms all other configurations.

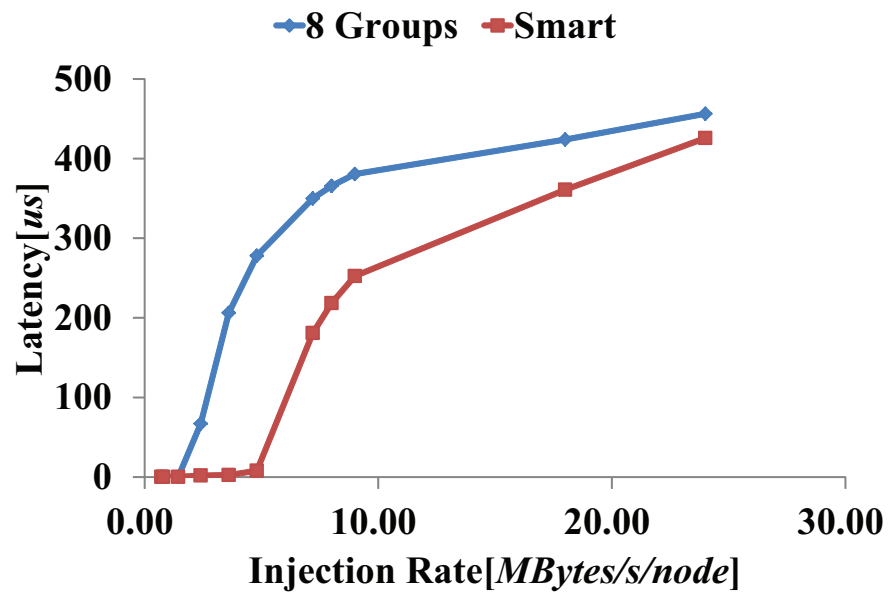
6.6 FORNoC under Partially Localized and Localized Probabilistic Traffic Patterns

In opposition to the uniform random traffic pattern in which communication is uniformly distributed throughout the network, we also evaluate the performance of FORNoC for partially and localized traffic patterns.

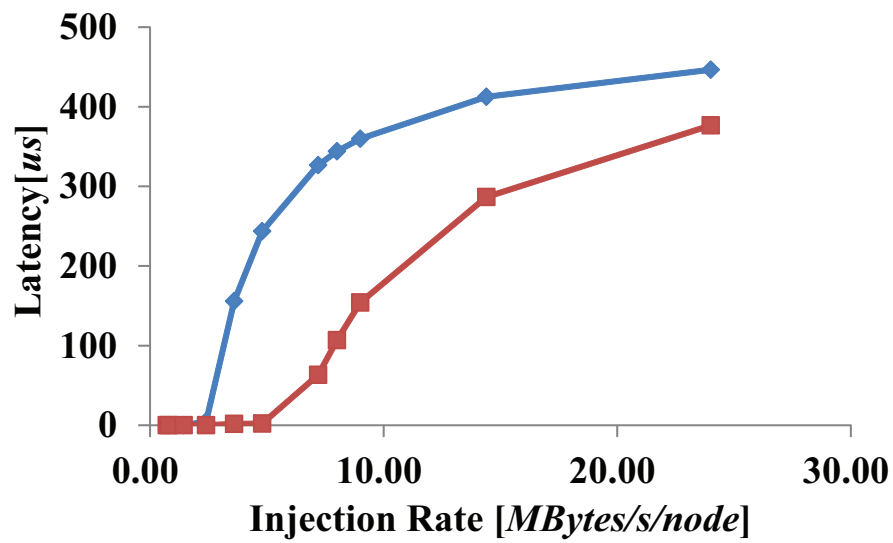
(1) Partially localized traffic pattern

We implement a neighbor communication pattern in which nodes communicate with their neighboring left and right node in a random manner. Figure 6.6 shows the performance in terms of latency (a), and bandwidth (b) for 8 Groups and smart selection mechanisms. As stated in Section 5.3, the smart selection further improves the performance of FORNoC by adapting the communication bandwidth to the network utilization.

For localized traffic, we use a hotspot traffic pattern. A node is randomly chosen as hotspot node, and all other nodes communicate with that node. Figure 6.7 shows the performance in terms of latency for SP (a) and DP (b) message injections, respectively for 8 groups and smart allocation mechanisms. The smart selection mechanisms outperform the 8 groups allocation technique.



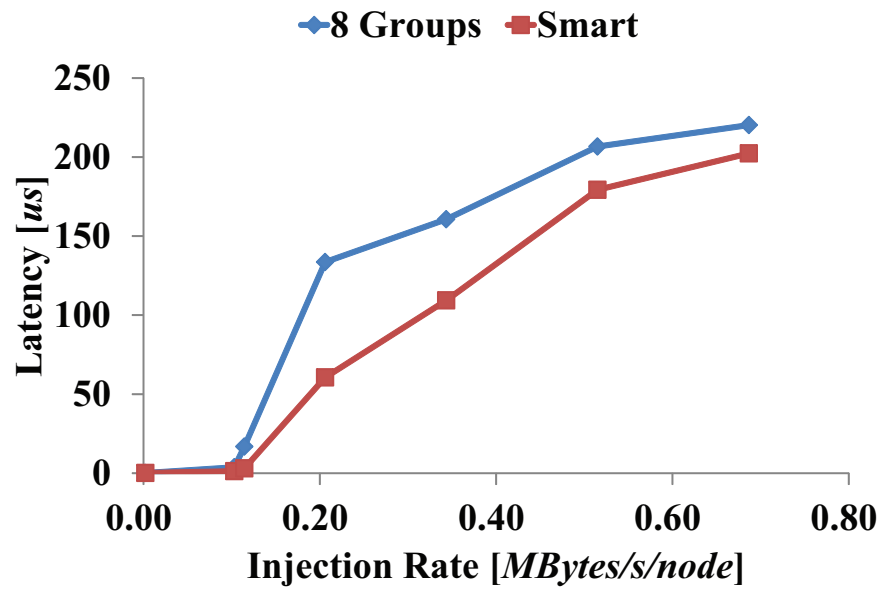
a) SP injection



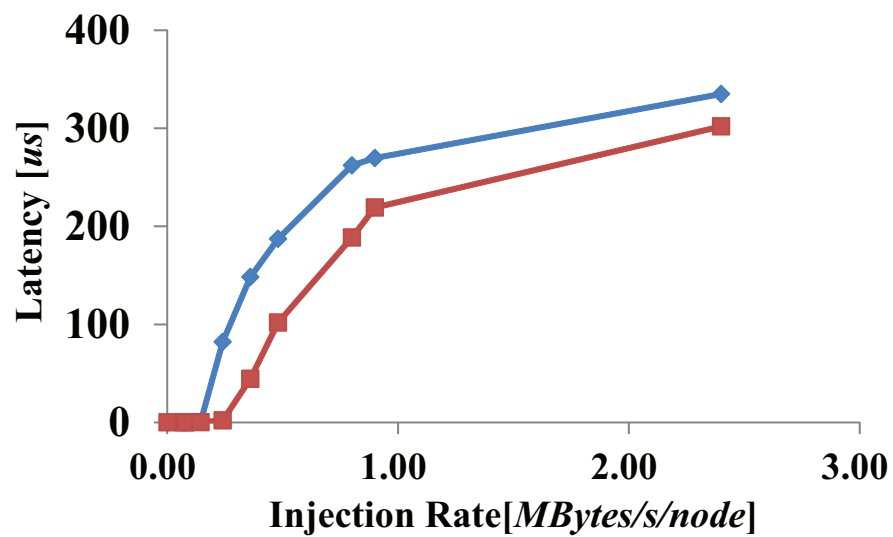
b) DP injection

Figure 6.6: Neighbor Traffic pattern.

(2) Localized traffic pattern



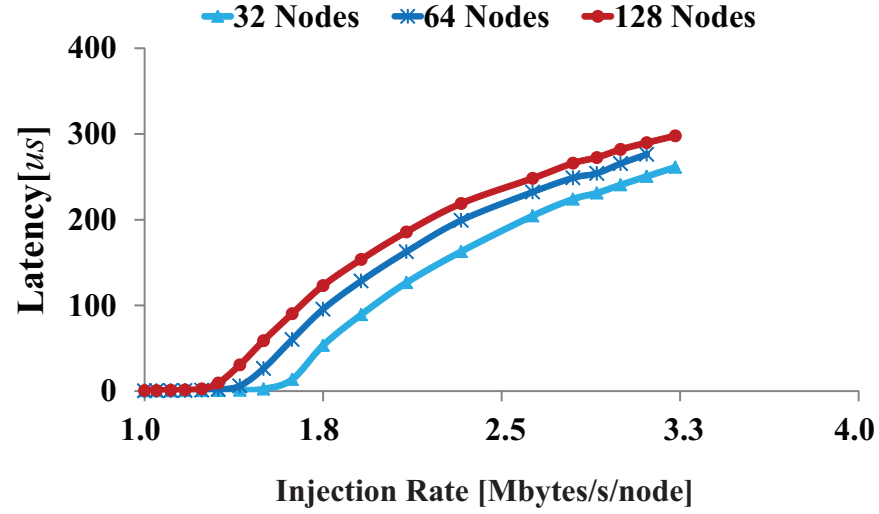
a) SP injection



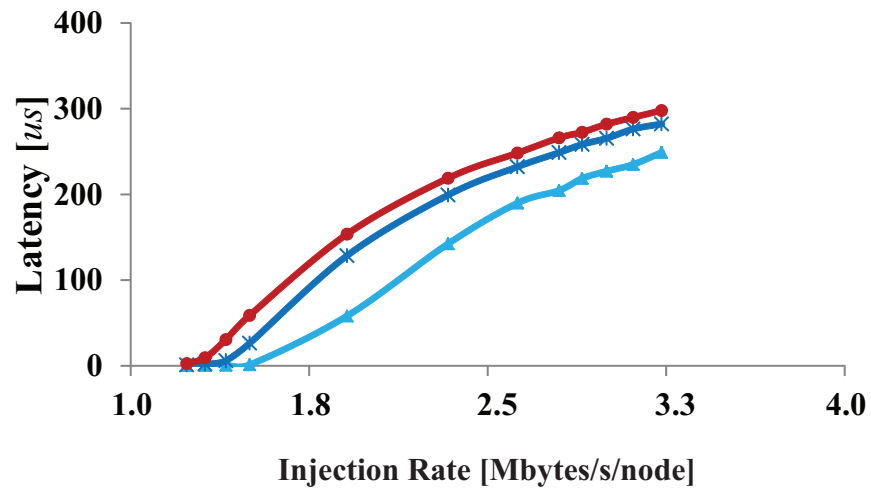
b) DP injection

Figure 6.7: Hotspot traffic pattern.

6.7 Scalability



a) Baseline Selection



b) Contention Based Selection

Figure 6.8: Average latencies for 32, 64, and 128-node under uniform random traffic.

Figure 6.8 shows the latency versus the injection rate for 32, 64, and 128-node networks under uniform traffic pattern for Baseline, and congestion based selection mechanisms. Although for larger networks: i) average distance is longer (more network nodes); ii) network saturates with smaller load because disjoint paths on the ring are reduced, the results show that the performance of FORNoC is scalable for both selection mechanisms.

6.8 Conclusion

In this chapter, we have proposed scalable photonic NoC architecture, which combines static and dynamic wavelength allocation communication mechanisms. The architecture takes advantage of both low-overhead/low-bandwidth of static, and high-overhead/high-bandwidth dynamic communications using wavelength allocation selection techniques, based on message size (normal selection, grouping), and congestion information (congestion based and smart selections).

Performance evaluation results under various probabilistic traffic patterns show that our proposed fully optical ring network (FORNoC) presents a good performance using adequate selection techniques. We also showed that our architecture reduces considerably the energy consumption necessary for arbitration compared to hybrid ring and mesh NoCs. A comparison with other previous work in term of architecture hardware cost shows that our architecture can be an attractive cost-performance efficient interconnection infrastructure for future SoCs and CMPs.

CHAPTER 7

SUMMARY

Silicon photonics Network-on-Chips (NoCs) have emerged as an attractive solution to alleviate the high power consumption of traditional electrical interconnects. Future NoC designs need to take full advantage of their advance to achieve high performance and low energy consumption communication infrastructure for future CMPs and SoCs. This Chapter summarizes our proposals described in the thesis and highlights some future works.

7.1 Conclusion

In this work, we propose three methods to take advantage of today's state-of-the-art on-chip interconnects. Firstly a low latency path setup network is proposed for hybrid planar NoCs using predictive switching and path reservation techniques. Second, we propose an electrical crossbar optical ring hybrid architecture which further improves the performance of hybrid electrical-optical interconnects. And finally, a fully optical ring NoC that combines static and dynamic wavelength allo-

cation communication mechanisms is presented. A different wavelength-channel is statically allocated to each destination node for light weight communication. Contention of simultaneous communication requests from multiple source nodes to the destination is solved by a token based arbitration for the particular wavelength-channel. For heavy load communication, a multiwavelength-channel is available by requesting it in execution time from source node to a special node that manages dynamic allocation of the shared multiwavelength-channel among all nodes. We combine these static and dynamic communication mechanisms in a same network that introduces selection techniques based on message size (baseline, Grouping), and congestion information (congestion based, and smart selections). Using a photonic NoC simulator based on Phoenixsim, we evaluate the architectures under uniform random, neighbor, and hotspot traffic patterns. Simulation results show that the fully optical ring NoC presents a good performance by utilizing adequate static and dynamic channels based on the selection techniques. We also show that the fully optical NoC architecture can reduce the energy consumption considerably compared to hybrid photonic ring and mesh NoCs. A comparison with several previous works in term of architecture hardware cost shows that our architecture can be an attractive cost-performance efficient interconnection infrastructure for future SoCs and CMPs.

7.2 Future works

An improvement to this work is to investigate fault tolerance ability for our architecture. As we use a single manager node which allocates path for the dynamic communication, when the node is faulty all dynamic communication wavelengths will become unavailable. Another improvement to this work is to analyze FORNoC behavior using real application communication traffic patterns.

Publications

Journal Papers (with reviews):

(1) Cisse Ahmadou Dit ADI, Michihiro Koibuchi, Masato Yoshimi Hidetsugu Irie, Tsutomu Yoshinaga: "A Fully Optical Ring Network-on-Chip with Static and Dynamic Wavelength Allocation", IEICE Transaction on Information and Systems Vol.E96-D, No.12, Dec. 2013. "Accepted "; [Chap.5,6].

(2) Cisse Ahmadou Dit ADI, Hiroki Matsutani, Michihiro Koibuchi, Hidetsugu Irie, Takefumi Miyoshi, Tsutomu Yoshinaga: "An Efficient Path Setup for a Hybrid Photonic Network-on-Chip." The International Journal of Networking and Computing (IJNC) 1(2): pp. 244-259 (2011); [Chap.3].

Proceedings of International Conferences (with reviews):

(3) Cisse Ahmadou Dit ADI, Ping Qiu, Hidetsugu Irie, Takefumi Miyoshi, Tsutomu Yoshinaga "OREX: An Optical Ring with Electrical Crossbar Hybrid Photonic Network-on-Chip". Proceedings of International Workshop on Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA 2010); [Chap.4,5,6].

(4) Cisse Ahmadou Dit ADI, Hiroki Matsutani, Michihiro Koibuchi, Hidetsugu Irie, Takefumi Miyoshi, Tsutomu Yoshinaga: "An Efficient Path Setup for a Photonic Network-on-Chip." Workshop in conjunction with IEEE ICNC 2010, Proceedings of the 2nd Workshop on Ultra Performance and Dependable Acceleration Systems (UPDAS'10): pp. 156-161; [Chap.3].

Other Publications:

Without reviews:

(5) Ping Qiu, Cisse Ahmadou Dit ADI, Hidetsugu Irie, Tsutomu Yoshinaga: "A Token-based Fully Photonic Network-on-Chip with Dynamic Wavelength Allocation". Proceedings of the International Workshop on Modern Science and Technology (IWMST 2012): pp. 39-44; [Chap.5,6].

With reviews:

(6) Yicheng Guan, Cisse Ahmadou Dit ADI, Takefumi Miyoshi, Michihiro Koibuchi, Hidetsugu Irie, and Tsutomu Yoshinaga: "Throttling Control for Bufferless Routing in On-Chip Networks" IEEE CS Proceedings of 6th IEEE International Symposium on Embedded Multicore SoCs (MCSoc-12): pp. 37-44.

BIBLIOGRAPHY

- [1] A. Biberman, K. Preston, G. Hendry, N. Sherwood-Droz, J. Chan, J. S. Levy, M. Lipson, and K. Bergman. Photonic Network-on-Chip Architectures Using Multilayer Deposited Silicon Materials for High-Performance Chip Multiprocessors. *ACM Journal on Emerging Technologies in Computing Systems*, 7(2):7:1–7:25, June 2011.
- [2] M. Briere, B. Girodias, Y. Bouchebaba, G. Nicolescu, F. Mieyeville, F. Gaffiot, and I. O'Connor. System Level Assessment of an Optical NoC in an MPSoC Platform. In *Proceedings of Design Automation Test in Europe (DATE)*, pages 1–6, april 2007.
- [3] J. Chan, G. Hendry, A. Biberman, K. Bergman, and L. Carloni. Phoenixsim: A Simulator for Physical-Layer Analysis of Chip-Scale Photonic Interconnection Networks. In *Proceedings of Design Automation Test in Europe (DATE)*, pages 691 –696, march 2010.
- [4] A. D. A. Cisse, P. Qiu, H. Irie, T. Miyoshi, and Y. T. OREX: A Hybrid Photonic Network-on-Chip of Optical Ring and Electrical Crossbar. In *Proceedings of the*

International Workshop on Innovative Architecture for Future Generation High-Performance Processors and Systems (IWIA), 2010.

- [5] M. Dahlem, M. Popovic, C. Holzwarth, A. Khilo, T. Barwicz, H. Smith, F. Kartner, and E. Ippen. Electronic-Photonic Integrated Circuits in Silicon-on-Insulator Platforms. In *Proceedings of the XXXth General Assembly and Scientific Symposium (URSI)*, page 1, aug. 2011.
- [6] W. Dally and B. Towles. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003.
- [7] H. Gilbert, C. Johnnie, K. Shoaib, O. Lenny, S. John, P. C. Luca, and B. Keren. Silicon Nanophotonic Network-on-Chip Using TDM Arbitration. In *Proceedings of the 18th IEEE Symposium on High Performance Interconnects (HOTI)*, pages 88–95, Los Alamitos, CA, USA, 2010. IEEE Computer Society.
- [8] M. Haurylau, G. Chen, H. Chen, J. Zhang, N. A. Nelson, D. H. Albonesi, E. G. Friedman, and P. M. Fauchet. On-Chip Optical Interconnect Roadmap: Challenges and Critical Directions. *IEEE Journal of Selected Topics in Quantum Electronics*, 12(6):1699–1705, Nov.-dec. 2006.
- [9] C. W. Holzwarth, J. S. Orcutt, L. Hanqing, M. A. Popovic, V. Stojanovic, J. L. Hoyt, R. J. Ram, and H. I. Smith. Localized Substrate Removal Technique Enabling Strong-Confinement Microphotonics in Bulk Si cmos Processes. In *Proceedings of the Conference on Quantum Electronics and Laser Science (CLEO/QELS)*, pages 1 –2, may 2008.
- [10] P. Jacquet, W. Szpankowski, and I. Apostol. A Universal Predictor based on Pattern Matching. *IEEE Transactions on Information Theory*, 48(6):1462 –1472, June 2002.

- [11] C. C. Kenneth, R. R. Miriam, A. B. Bruce, M. B. Audrey, L. K. David, and D. Paul. Challenges for On-Chip Optical Interconnects. In *Proceedings of SPIE on Optoelectronic Integration on Silicon II*, pages 133–143, 2005.
- [12] N. Kirman, M. Kirman, R. K. Dokania, J. F. Martinez, A. B. Apsel, M. A. Watkins, and D. H. Albonesi. Leveraging Optical Technology in Future Bus-based Chip Multiprocessors. In *Proceedings of the 39th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO 39*, pages 492–503, Washington, DC, USA, 2006. IEEE Computer Society.
- [13] S. Koohi and S. Hessabi. Contention-Free On-chip Routing of Optical Packets. In *Proceedings of the 3rd ACM/IEEE International Symposium on Networks-on-Chip (NOCS)*, pages 134 –143, may 2009.
- [14] C. Li, M. Browning, P. V. Gratz, and S. Palermo. LumiNOC: A Power-Efficient, High-Performance, Photonic Network-on-Chip for Future Parallel Architectures. In *Proceedings of the 21st International Conference on Parallel Architectures and Compilation Techniques (PACT)*, PACT '12, pages 421–422, New York, NY, USA, 2012. ACM.
- [15] M. Lipson. Guiding, Modulating, and Emitting Light on Silicon-Challenges and Opportunities. *Journal of Lightwave Technology*, 23(12):4222–4238, Dec 2005.
- [16] H. Matsutani, M. Koibuchi, H. Amano, and T. Yoshinaga. Prediction Router: Yet Another Low Latency On-Chip Router Architecture. In *Proceedings of the 15th IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pages 367 –378, feb 2009.
- [17] H. Matsutani, M. Koibuchi, H. Amano, and T. Yoshinaga. Prediction Router: A Low-Latency On-Chip Router Architecture with Multiple Predictors. *IEEE Transactions on Computer*, 60(6):783 –799, june 2011.

- [18] R. Morris and A. Kodi. Exploring the Design of 64- and 256-core Power Efficient Nanophotonic Interconnect. *IEEE Journal of Selected Topics in Quantum Electronics*, 16(5):1386 –1393, sept.-oct. 2010.
- [19] Y. Pan, J. Kim, and G. Memik. FlexiShare: Channel Sharing for an Energy-Efficient Nanophotonic Crossbar. In *Proceedings of the 16th IEEE International Symposium on High Performance Computer Architecture (HPCA)*, pages 1–12, 2010.
- [20] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang, and A. Choudhary. Firefly: Illuminating Future Network-on-Chip with Nanophotonics. In *Proceedings of the 36th International Symposium on Computer Architecture (ISCA)*, ISCA '09, pages 429–440, New York, NY, USA, 2009. ACM.
- [21] X. Qianfan, M. Sasikanth, S. Brad, S. Jagat, and L. Michal. 12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators. *Journal of Optics Express*, 15(2):430–436, Jan 2007.
- [22] M. Reshotko, B. A. Block, J. Ben, and P. Chang. Waveguide Coupled Ge-on-oxide Photodetectors for Integrated Optical Links. In *Proceedings of the 5th IEEE International Conference on Photonics, Group IV*, pages 182 –184, sept. 2008.
- [23] A. Shacham, K. Bergman, and L. Carloni. The Case for Low-Power Photonic Networks on Chip. In *Proceedings of the 44th IEEE/ACM Design Automation Conference (DAC)*, pages 132 –135, 2007.
- [24] A. Shacham, K. Bergman, and L. P. Carloni. On the Design of a Photonic Network-on-Chip. In *Proceedings of the International Symposium on Networks-on-Chip (NOCS)*, pages 53 –64, May 2007.

- [25] A. Shacham, K. Bergman, and L. P. Carloni. Photonic Networks-on-Chip for Future Generations of Chip Multiprocessors. *IEEE Transactions on Computers*, 57(9):1246–1260, sept. 2008.
- [26] A. Shacham, B. G. Lee, A. Biberman, K. Bergman, and L. P. Carloni. Photonic NoC for DMA Communications in Chip Multiprocessors. In *Proceedings of the 15th IEEE Symposium on High-Performance Interconnects (HOTI)*, HOTI '07, pages 29–38, Washington, DC, USA, 2007. IEEE Computer Society.
- [27] K. Somayyeh and H. Shaahin. All-Optical Wavelength-Routed Architecture for a Power-Efficient Network on Chip. *IEEE Transactions on Computers*, 99(PrePrints), 2012.
- [28] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. Beausoleil, and J. Ahn. Corona: System Implications of Emerging Nanophotonic Technology. In *Proceedings of the 35th International Symposium on Computer Architecture (ISCA)*, pages 153–164, june 2008.
- [29] I. Young, E. Mohammed, J. Liao, A. Kern, S. Palermo, B. Block, M. Reshotko, and P. Chang. Optical I/O Technology for Tera-Scale Computing. In *Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC)*, pages 468–469, 469a, feb. 2009.
- [30] A. V. Yurii and J. M. Sharee. Losses in single-mode silicon-on-insulator strip waveguides and bends. *Journal of Optics Express*, 12(8):1622–1631, 2004.